FERMILAB-PUB-15-309-AE

The DES Science Verification Weak Lensing Shear Catalogs

M. Jarvis[†1], E. Sheldon[2], J. Zuntz[3], T. Kacprzak[4], S. L. Bridle[3], A. Amara[4], R. Armstrong[5], M. R. Becker[6,7], G. M. Bernstein[1], C. Bonnett[8], C. Chang[4], R. Das[9], J. P. Dietrich[10,11], A. Drlica-Wagner[12], T. F. Eifler[1,13], C. Gangkofner[10,14], D. Gruen[15,11], M. Hirsch[16], E. M. Huff[17,18], B. Jain[1], S. Kent[12], D. Kirk[16], N. MacCrann[3], P. Melchior[17,18], A. A. Plazas[13], A. Refregier[4], B. Rowe[16], E. S. Rykoff[7,19], S. Samuroff[3], C. Sánchez[8], E. Suchyta[17,18], M. A. Troxel[3], V. Vikram[20], T. Abbott[21], F. B. Abdalla[16], S. Allam[12], J. Annis[12], A. Benoit-Lévy[16], E. Bertin[22,23], D. Brooks[16], E. Buckley-Geer[12], D. L. Burke[7,19], D. Capozzi[24], A. Carnero Rosell[25,26], M. Carrasco Kind[27,28], J. Carretero[29,8], F. J. Castander[29], M. Crocce[29], C. E. Cunha[7], C. B. D'Andrea[24], L. N. da Costa[25,26], D. L. DePoy[30], S. Desai[10,14], H. T. Diehl[12], P. Doel[16], A. Fausti Neto[25], B. Flaugher[12], P. Fosalba[29], J. Frieman[12,31], E. Gaztanaga[29], D. W. Gerdes[9], R. A. Gruendl[27,28], G. Gutierrez[12], K. Honscheid[17,18], D. J. James[21], K. Kuehn[32], N. Kuropatkin[12], O. Lahav[16], T. S. Li[30], M. Lima[33,25], M. March[1], P. Martini[17,34], R. Miquel[35,8], J. J. Mohr[10,14,15], E. Neilsen[12], B. Nord[12], R. Ogando[25,26], K. Reil[19], A. K. Romer[36], A. Roodman[7,19], M. Sako[1], E. Sanchez[37], V. Scarpine[12], M. Schubnell[9], I. Sevilla-Noarbe[37,27], R. C. Smith[21], M. Soares-Santos[12], F. Sobreira[12,25], M. E. C. Swanson[28], G. Tarle[9], J. Thaler[38], D. Thomas[24], A. R. Walker[21], R. H. Wechsler[6,7,19]

(The DES Collaboration)

Affiliations are listed at the end of the paper

arXiv:1507.05603v2 [astro-ph.IM] 22 Jul 2015

ABSTRACT

We present weak lensing shear catalogs for 139 square degrees of data taken during the Science Verification (SV) time for the new Dark Energy Camera (DECam) being used for the Dark Energy Survey (DES). We describe our object selection, point spread function estimation and shear measurement procedures using two independent shear pipelines, IM3SHAPE and NGMIX, which produce catalogs of 2.12 million and 3.44 million galaxies respectively. We detail a set of null tests for the shear measurements and find that they pass the requirements for systematic errors at the level necessary for weak lensing science applications using the SV data. We also discuss some of the planned algorithmic improvements that will be necessary to produce sufficiently accurate shear catalogs for the full 5-year DES, which is expected to cover 5000 square degrees.

Contents

1 Introduction **2**

2 Data **4**
- 2.1 Object Catalog 5
- 2.2 Galaxy Selection 6
- 2.3 Astrometry 6

3 Requirements on Systematic Errors **7**
- 3.1 Shear Correlation Functions 8
- 3.2 Multiplicative and Additive Errors . . 8
- 3.3 PSF Leakage 9
- 3.4 PSF Model Errors 10

4 PSF Estimation **11**
- 4.1 Initial Identification of Stars 11
- 4.2 Selection of PSF Stars 12
- 4.3 PSF Measurement and Interpolation . 13
- 4.4 PSF Model Diagnostics 14

5 Multi-Epoch Data Structures **15**
- 5.1 Exposure Selection 16

[†]Corresponding author; mjarvis@physics.upenn.edu

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

5.2 Masks 16

6 Simulations **17**
6.1 GREAT-DES 17
6.2 End-to-end Simulation 19

7 Shear Measurement **19**
7.1 Multi-epoch Fitting 20
7.2 Signal-to-noise Ratio 20
7.3 Shear Measurements with IM3SHAPE 21
7.3.1 Bulge or disk model 22
7.3.2 Shear calibration 22
7.3.3 Diagnostics 24
7.3.4 Galaxy weights 24
7.4 Shear Measurements with NGMIX . . 24
7.4.1 Exponential disk model . . . 24
7.4.2 Image fitting 25
7.4.3 Shear estimation 25
7.4.4 Ellipticity prior 26
7.5 Blinding 26

8 Tests of the Shear Measurements **27**
8.1 Spatial Tests 28
8.1.1 Position in the field of view . 28
8.1.2 Position on CCD 29
8.1.3 Tangential shear around field centers 29
8.2 PSF Tests 30
8.2.1 PSF leakage 30
8.2.2 Star/galaxy cross-correlation . 31
8.2.3 Tangential shear around stars . 32
8.3 Galaxy Property Tests 32
8.3.1 Galaxy signal-to-noise 32
8.3.2 Galaxy size 33
8.4 B-mode Statistics 34
8.5 Calibration Tests 34
8.6 Cross-catalog Comparisons 35
8.6.1 Tangential shear ratio 36
8.6.2 Differential shear correlations 38
8.7 Summary of Systematics Tests 39

9 Shear Catalogs **40**
9.1 Final Galaxy Selection 40
9.2 Applying the Calibration/Sensitivity . 41
9.3 Effective Number Density 42

10 Summary and Discussion **43**

A Multi-Epoch Data Structures **46**
A.1 Bulding the MEDS Files 46
A.2 Structure of the MEDS Files 46

B Catalog Flags **48**
B.1 IM3SHAPE Flags 48
B.2 NGMIX Flags 50

C Details of the Shear Catalogs **50**

1. Introduction

Weak gravitational lensing provides a powerful statistical tool for studying the distribution of mass in the Universe. Light traveling from distant galaxies to Earth is deflected by the gravitational field of mass concentrations along the path. This deflection distorts the observed light distribution of galaxies, and when this distortion is very small, stretching the surface brightness profile by of order a few percent or less, it is referred to as "weak lensing".

The weak lensing distortion includes both a stretching component called "shear" and a dilation component called "convergence". Here we focus on the shear. The observed shear field can be used to make maps of the matter in the universe, uncover the mass profiles of galaxies and clusters of galaxies, and even test theoretical models of dark energy.

In order to reach its full potential as a probe of dark matter and dark energy, shear measurement must be extremely accurate. Each galaxy is typically stretched by about 2%, whereas the intrinsic unknown ellipticity of the galaxy before being lensed is an order of magnitude larger. This "shape noise" constitutes the primary statistical uncertainty for weak lensing measurements. Nonetheless, by measuring the shapes of millions of galaxies, the Dark Energy Survey (DES) and other current surveys can expect to make precise measurements of the mean shear with fractional statistical uncertainties as low as 1%. Future surveys may reach 0.1%. This implies that systematic errors (i.e. biases) in the shape measurements need to be controlled at a level approximately 3 orders of magnitude smaller than the shape noise on each measurement.

There are many potential sources of systematic error that can bias the shape measurements used for estimating shears. The galaxy images are blurred and smeared when the photons pass through the atmo-

sphere, the telescope optics, and the detector, leading to a spatially and temporally variable point-spread function (PSF). The images are stretched by distortion from the telescope and sometimes by features of the detector. The images are pixellated and have various sources of noise. Detector defects, cosmic rays, satellite trails, and other artifacts in the data can lead to some pixels not being used, and measurement algorithms must deal properly with this "missing data". Flux from nearby galaxies or stars can obfuscate the determination of the observed intensity profile. All of these phenomena must be included in the analysis at very high accuracy if systematic uncertainties are to be sub-dominant to statistical uncertainties.

Previous studies have taken a range of approaches to measuring galaxy shapes, typically falling into one of two categories. Moments-based methods (e.g. Kaiser et al. 1995; Rhodes et al. 2000; Melchior et al. 2011) involve measuring second and higher-order moments of the galaxy and the PSF. Model-fitting methods (e.g. Massey & Refregier 2005; Nakajima & Bernstein 2007; Miller et al. 2013) involve fitting a PSF-convolved galaxy model to the data. A number of blind challenges of shear measurements have been carried out to assess progress in a uniform way across the international shear measurement community: the Shear TEsting Programme (STEP Heymans et al. 2006; Massey et al. 2007) and the GRavitational lEnsing Accuracy Testing (GREAT) Challenges (Bridle et al. 2009, 2010; Kitching et al. 2010, 2012; Mandelbaum et al. 2014, 2015). The wide variety of shear measurement methods and their performance on these benchmarks are summarized there. The two shear algorithms presented in this work, IM3SHAPE (Zuntz et al. 2013), and NGMIX, are both of the model-fitting variety (cf. §7).

Most shear measurement methods are biased in the low signal-to-noise (S/N) regime, where the impact of pixel noise on the shape measurement of each galaxy becomes significant. This "noise bias" effect was first discussed in Bernstein & Jarvis (2002) and Hirata et al. (2004), and was found to be the most significant of the effects studied in the GREAT08 Challenge (Bridle et al. 2009, 2010). It was derived analytically for maximum-likelihood methods in Refregier et al. (2012), in the context of direct estimation in Melchior & Viola (2012), and quantified in the context of future surveys in Kacprzak et al. (2012).

Complex galaxy morphologies can also bias shear measurements (Massey et al. 2007; Lewis 2009; Voigt & Bridle 2010; Bernstein 2010; Melchior et al. 2010; Zhang & Komatsu 2011). This "model bias" can arise even for simple galaxy profiles if the model being used does not match reality. Model bias was found to be around 1% for bulge+disk model fitting methods, and the interplay with noise bias was found to be small (Kacprzak et al. 2014). The GREAT3 challenge (Mandelbaum et al. 2014, 2015) included realistic galaxy morphologies, and those authors found that the mean model bias was $\sim 1\%$ for a wide range of methods. The Fourier Domain Nulling approach (Bernstein 2010) provides a potential solution to this problem, which may be able to avoid model bias altogether.

One strategy to account for these biases is to apply a multiplicative correction factor calibrated from image simulations. This can take the form of a single constant bias correction applied to all galaxies (e.g. Schrabback et al. 2007), or it can vary according to galaxy properties such as the signal-to-noise ratio (Schrabback et al. 2010; Gruen et al. 2013) and size (von der Linden et al. 2014). For the IM3SHAPE shear measurements, we calibrate biases as a function of both of these parameters, as done by Kacprzak et al. (2012). A significant improvement in the current analysis lies in our modeling of additive systematic errors as proportional to PSF ellipticity, which we also apply as a calibration (cf. §7.3.2).

A different strategy to account for noise bias (although not model bias) is to include the known distribution of intrinsic galaxy shapes as Bayesian prior information and fully sample the posterior likelihood surface. Miller et al. (2007) proposed a first order approximation to this, and a more rigorous treatment was given by Bernstein & Armstrong (2014). For the NGMIX shear measurements, we follow the approach of Miller et al. (2007) (cf. §7.4.3).

Each part of the sky in a weak lensing survey is generally observed multiple times. Most commonly, the shape measurements are made on coadded images of these multiple exposures (e.g. Wittman et al. 2000; Van Waerbeke et al. 2000; Heymans et al. 2005; Leauthaud et al. 2007; Fu et al. 2008). While coadded images reduce the total data volume, making data handling easier, differences in the PSFs between the epochs complicate the modeling of the coadded PSF and often introduce spurious effects that are problematic for the most sensitive shear probes. Multi-epoch methods (Tyson et al. 2008; Bosch 2011; Miller et al. 2013) instead simultaneously use all individual exposures of a galaxy with the corresponding single-epoch PSF models and

weights, thereby avoiding these problems.

The current state-of-the-art weak lensing shear measurement comes from the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012b), which observed 154 square degrees of sky and measured 7.6 million galaxy shapes. They discovered that the previous CFHTLenS analysis (Fu et al. 2008), using coadd images, had significant systematic errors and that switching to a multi-epoch method (Miller et al. 2013) was superior. We use similar multi-epoch algorithms in this work (cf. §7.1).

For removing problematic data, the CFHTLenS analysis trimmed the survey area to only those fields in which the shape catalogs passed certain systematic tests. We use a somewhat different strategy in our analysis. We blacklist single-epoch images that fail tests of the image quality, the astrometric solution, or the PSF model, and exclude them from the multi-epoch fitting process (cf. §5.1).

In this paper we present the shear catalog for the DES Science Verification (SV) data, described in §2. We derive requirements for our systematic uncertainties in §3. The PSF model is described and tested in §4. To facilitate multi-epoch shear measurements, we developed Multi-Epoch Data Structures (MEDS), which we describe in §5. Two sets of simulations that we used for calibration and testing are presented in §6. We present our two shear estimation codes, IM3SHAPE and NGMIX in §7. Then we submit our catalogs to a suite of null tests, described in §8, which constitutes the main results of this paper. Finally, we describe our final shear catalogs in §9 and conclude in §10. Appendices provide more information on the data structures and catalog flags. A flowchart outlining the main stages in the production of the shear catalogs is shown in Figure 1.

2. Data

The Dark Energy Camera (DECam; Flaugher et al. 2015; Diehl 2012; Honscheid et al. 2012) was installed on the 4m Victor M. Blanco Telescope at the Cerro Tololo Inter-American Observatory (CTIO) in Chile from June, 2011, to September, 2012 (Diehl et al. 2014). The first light ceremony was September 12, 2012.

DECam holds sixty-two 2048×4096 science CCDs, four 2048×2048 guider CCDs, and eight 2048×2048 focus and alignment CCDs, for a total of 570 megapixels covering a roughly hexagonal footprint. The CCDs were fabricated at Teledyne Dalsa[1], further processed by Lawrence Berkeley National Laboratory (LBNL), and assembled and tested at Fermilab[2]. Each CCD is 250 microns thick and fully depleted, with two amplifiers per CCD.

The DECam field of view has a diameter of 2.2 degrees on the sky. Unfortunately, one the 62 science CCDs was damaged during commissioning, so we have only 61 working CCDs[3]. The total usable footprint of an exposure, excluding the gaps between the CCDs, totals 2.6 square degrees. Five filters are used during normal survey operations, g, r, i, z, Y, exchanged using an automated shutter-filter system (Tarlé et al. 2010).

The Dark Energy Survey (DES) officially started taking survey data in August, 2013 (Diehl et al. 2014). It will cover about 5000 square degrees in the South Galactic Cap region, with $\sim$10 visits per field in the g, r, i and z bands (two visits per year), for a 10σ limiting magnitude of about 24.1 in the i-band. In addition to the main survey, the DES supernova survey contains smaller patches optimized for time-domain science, which are visited more often, and which are useful as a deeper dataset observed with the same instrument.

Before the start of the main survey, a small Science Verification (SV) survey was conducted from November 2012 to February 2013. The strategy was to observe the SV area at 10 different epochs, mimicking the number of visits and total image depth planned for the full 5-year DES survey. Significant depth variations exist in the SV data due to weather, issues with the telescope, and no data quality checks to ensure uniformity (cf. Leistedt et al. 2015). For the current study we restricted our measurements to the largest portion of the SV area, known as SPT-East (SPT-E for short), an area of approximately 139 square degrees contained within the eastern part of the region observed by the South Pole Telescope (SPT; Carlstrom et al. 2011).

The SV data are reduced by the DES Data Management (DESDM) system (Mohr et al. 2012; Desai et al. 2012), resulting in calibrated and background-subtracted images. Catalogs are produced using the software package Source Extractor (SEXTRACTOR;

[1] https://www.teledynedalsa.com

[2] Fermilab is operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

[3] One additional science CCD failed in the first year of the DES main survey, but it was still functional for the work presented here.

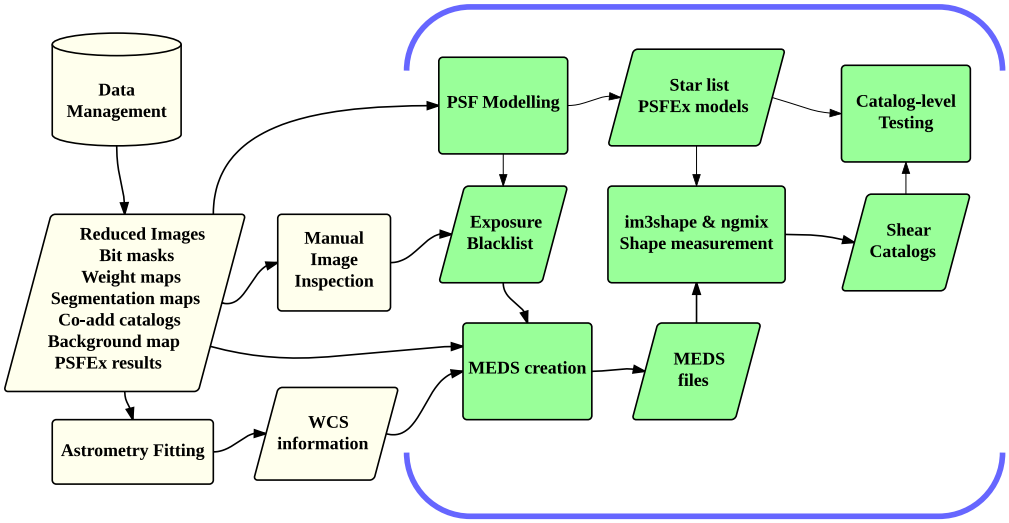

Fig. 1.— A flowchart showing the main stages in the production of the shear catalogs. The items inside the blue bracket are done by the weak lensing group in DES and are the principal subject of this paper.

Bertin & Arnouts 1996; Bertin 2011). The point spread function is characterized using the PSFEx package (Bertin 2011; for more details, see §4).

On a set of pre-defined areas of sky, all overlapping single-epoch images are registered and combined into a coadd image using the SCAMP and SWARP packages (Bertin et al. 2002; Bertin 2006).

For weak lensing we use these coadd images only for object detection, deblending, fluxes (for use in photometric redshift measurements, see Sánchez et al. 2014; Bonnett et al. 2015), and for the detailed informational flags which are important for determining a good set of galaxies to use for shear measurement.

In contrast to previous work on DES data by Melchior et al. (2015), we perform object shape measurement directly on all available single-epoch images in which an object was observed, using multi-epoch fitting techniques. See §5 for more details of how we repackage the data for multi-epoch fitting and §7.1 for a description of the multi-epoch measurement process.

2.1. Object Catalog

The starting point for our object catalog is the "SVA1 Gold Catalog" (Rykoff et al. 2015), which excludes regions of the data that are known to be problematic in some way: due to imaging artifacts, scattered light, failed observations, etc. The selection criteria for the Gold Catalog include the following:

- Require object to have been observed at least once in each of the g, r, i, and z bands.
- Require Declination to be north of $61°S$ to avoid the Large Magellanic Cloud and R Doradus, where the photometric calibration was found to have severe problems.
- Remove regions with a high density of objects with "crazy colors", i.e. those with any of the following: $g-r<-1$, $g-r>4$, $i-z<-1$, or $i-z>4$. Such regions are usually due to satellite trails, ghosts, scattered light, etc.
- Remove regions with a density less than 3σ below the mean density.
- Remove regions near bright stars. We eliminate a circular region around all stars detected in the 2-micron All Sky Survey (2MASS; Skrutskie et al. 2006) brighter than $J_M = 12$ with a mask radius of $r = (-10J_M + 150)$ arcseconds up to a maximum radius of 120 arcseconds.
- Remove regions with a concentration of objects with large centroid shifts between bandpasses. Some of these objects are just dropout galaxies or large galaxies with complex, wavelength-dependent substructure, but many are due to scattered light, ghosts, satellite trails, etc. 25% of such objects fall into 4% of the total area, so we remove all objects in that 4% on the assumption that the other nearby objects probably have corrupted shapes and photometry.

The full SPT-E area observed during SV totals 163 square degrees. Applying the above selection criteria

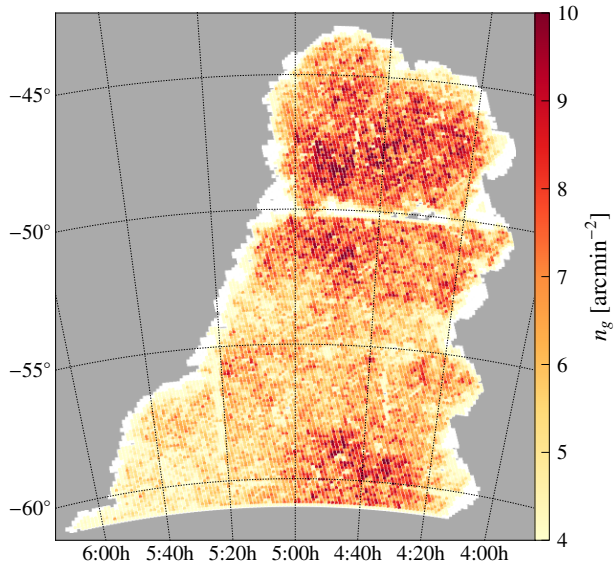

Fig. 2.— A HEALPIX map of the SPT-E region. The white background shows the full "Gold" area. The colors show the galaxy density in the NGMIX shear catalog. (The map for IM3SHAPE is qualitatively similar, although about 40% shallower.) The map has HEALPIX resolution `nside` = 512.

brings this down to 148 square degrees for the Gold Catalog.

The selection criteria listed above remove galaxies in a non-random way that varies across the sky. We characterize this selection using a geometrical "mask", implemented as a HEALPIX map (Górski et al. 2005). The HEALPIX map for the DES SPT-E region is shown in Figure 2. The white background represents the Gold Catalog area. The colored intensity represents the galaxy number density in the NGMIX catalog (cf. §7.4).

The region used for the weak lensing analysis is somewhat smaller than the full Gold Catalog region, because we additionally exclude CCD images with poor astrometric solutions (cf. §2.3), poor PSF solutions (cf. §4.2), and blacklisted CCDs containing bright stars, ghosts, airplanes etc. (cf. §5.1). The astrometric cuts in particular removed regions near the edge, since the solutions were poorly constrained there, resulting in a final area for the shear catalogs of 139 square degrees. The intensity map for the IM3SHAPE catalog (cf. §7.3) looks qualitatively similar, although it is about 40% shallower (cf. §9.3).

2.2. Galaxy Selection

The preliminary galaxy selection is performed using standard SEXTRACTOR outputs from the i-band detections in the Gold Catalog (Rykoff et al. 2015). The selection, in pseudo-code, is

```
bright_test = CLASS_STAR > 0.3
              AND MAG_AUTO < 18.0
locus_test = SPREAD_MODEL +
             3*SPREADERR_MODEL < 0.003
faint_psf_test = MAG_PSF > 30.0
                 AND MAG_AUTO < 21.0

galaxies = NOT bright_test
           AND NOT locus_test
           AND NOT faint_psf_test
```

Within DES, this is called the "Modest Classification" scheme. Bright stars are identified by the standard SEXTRACTOR classifier (`bright_test`). Fainter objects are considered stars if they are near the stellar locus in the `SPREAD_MODEL` measure introduced by Desai et al. (2012) (`locus_test`; see also Bouy et al. 2013). Objects that pass the `faint_psf_test` tend to be spurious detections and are considered "junk" in this classification. Our galaxy selection then includes every object not classified as either a star or junk, but many faint stars remain. Further selection criteria, described in §9.1 are applied to the catalogs based on measurements from the shear pipelines.

2.3. Astrometry

For each CCD image we must establish an astrometric solution, i.e. a map from pixel coordinates (x, y) to celestial coordinates (θ, ϕ), known as the World Coordinate System (WCS). Since the determination of galaxy shapes is done by a simultaneous fit to the pixel data for all single-epoch exposures covering the galaxy, any misregistration of the exposures will introduce spurious shear signals into the inferred galaxy shapes and sizes.

We found that the astrometric solutions provided by DESDM were not sufficiently accurate for our needs. They included misregistrations of more than 150 milliarcseconds on some CCDs, which induced unacceptably high systematic errors in the galaxy shapes. Here we describe the process we used to improve these solutions to the WCS.

Astrometric solutions for the SV exposures are assumed to take the form

$$P(\theta, \phi) = E\left(C(x, y)\right), \tag{2-1}$$

where P is a gnomonic projection from the (curved) sky onto a planar coordinate system, using a chosen field coordinate for the pole of the projection; E is an affine transformation chosen to be distinct for each CCD image of each exposure; and C is a cubic polynomial mapping that is common to all exposures in a given filter with a given CCD. In the nomenclature of the SCAMP code[4], C is the "instrument" solution, and E is the "exposure" solution.

The instrument solution C is derived as follows. We take a series of $\approx$ 20 exposures of a rich star field in succession, with the telescope displaced by angles ranging from 10 arcseconds up to the field of view of the camera. Coordinates of stars are determined in the pixel coordinates of each exposure, and we adjust the parameters of the map in equation 2-1 to minimize the internal disagreement between sky coordinates of all the observations of each star. The solution also minimizes the discrepancies between the positions of stars in the 2MASS (Skrutskie et al. 2006) catalog and our measurements of these stars, thereby anchoring the absolute pointing and scale of our astrometric maps.

All 20 coefficients of the cubic polynomial C are free for each of the 61 functional CCDs (cf. Diehl et al. 2014; Flaugher et al. 2015). While fitting the star field data, we force all CCDs in a given exposure to share a common affine map E, so there are 6 additional free parameters in the fit for each exposure. The instrument maps C derived in this way are assumed to apply to all SV exposures taken with the same CCD in the same filter. The process is repeated for each of the g, r, i, z, Y filters.

For each CCD image in the SV data, we determine an independent E function (6 degrees of freedom) in another round of fitting. This time we are minimizing the disagreements between positions reported for all CCDs that contribute to each DES coadd image. The coefficients of the affine transformations E are allowed to float, but the higher-order polynomials C are held fixed at the values determined from the star field data. These solutions again minimize residuals with respect to matching sources from the 2MASS catalog in order to fix the absolute position on the sky.

Note that the principal effects of differential chromatic refraction (DCR) are a shift and a shear along the direction toward zenith, which are both properly included as part of the affine transformation E for each CCD. We do not however make any attempt to address the intra-band chromatic effects related to DCR (cf. Plazas & Bernstein 2012; Meyers & Burchat 2015).

The RMS disagreement between sky positions of bright stars inferred from distinct DES exposures implies errors in the astrometric maps of 10–20 milliarcseconds RMS in each coordinate. These errors are seen to be coherent over arcminute scales in a given exposure, but are uncorrelated between distinct exposures. This suggests that the remaining relative astrometric errors are dominated by stochastic atmospheric distortions (cf. Heymans et al. 2012a). Indeed, equation 8 of Bouy et al. (2013) predicts an RMS astrometric residual due to the atmosphere of order 10 milliarcseconds for our field of view and exposure time.

There do remain some astrometric errors that are coherent over time and correlated with position on the detector array, and are well described as due to small components of the electric fields transverse to the surface of the CCD in some places (Plazas et al. 2014). These residuals are at the few milliarcsecond level, which is small enough to be irrelevant for SV data reductions.

3. Requirements on Systematic Errors

In this section, we derive the requirements for systematic uncertainties on the shear estimates for the DES SV data. These requirements will be used to assess the quality of the PSF and shear catalogs in subsequent sections.

Throughout this paper we will use the notation $e = e_1 + ie_2 = |e| \exp(2i\phi)$ as the complex-valued shape of each galaxy. We define the shape e such that the expectation value of the mean shape for an ensemble of galaxies is an estimate of the mean reduced gravitational shear acting on those galaxies

$$\langle e \rangle = g \equiv \frac{\gamma}{1-\kappa}, \tag{3-1}$$

where γ and κ are the shear and convergence, respectively (see e.g. Hoekstra 2013 for a review of weak lensing concepts and terminology).

For a galaxy with elliptical isophotes, one finds that $|e| = (a-b)/(a+b)$ satisfies equation 3-1, where a and b are the semi-major and semi-minor axes of the ellipse. However, galaxies do not in general have elliptical isophotes, so this definition is of little practical value. For the more general case, the estimator

$$e = \frac{I_{xx} - I_{yy} + 2iI_{xy}}{I_{xx} + I_{yy} + 2\sqrt{I_{xx}I_{yy} - I_{xy}^2}} \tag{3-2}$$

[4] http://www.astromatic.net/software/scamp

has been proposed by Seitz & Schneider (1997), where the second moments of the intensity profile $I(x, y)$ are defined as

$$I_{\mu\nu} = \frac{\int dxdyI(x,y)(\mu - \bar{\mu})(\nu - \bar{\nu})}{\int dxdyI(x,y)}. \quad (3\text{-}3)$$

But since neither shear algorithm in this paper uses equation 3-2 directly, we consider equation 3-1 to be the functional definition of what we mean by the shape of an arbitrary galaxy. See §7.3.1 and §7.4.1 for details about the IM3SHAPE and NGMIX estimators of e.

While equation 3-1 is our goal for the shape estimates in our catalog, it is inevitable that there will be systematic errors in the shape measurements. A convenient parameterization, based on one first proposed by Heymans et al. (2006), uses a first-order expansion of the form,

$$\langle e \rangle = (1 + m)g_{\rm true} + \alpha e_{\rm PSF} + c, \quad (3\text{-}4)$$

where $g_{\rm true}$ denotes the value that would be obtained from an ideal error-free shape estimator, m quantifies the *multiplicative error*, α measures the *leakage* of the PSF shape into the galaxy shapes, and c represents other sources of *additive error*.

Note that m can in principle be different for each of the two components e_1 and e_2. However, we find in practice that the two coefficients are generally very close to equal when they can be measured separately, so we simply take m to be a single real value here. Similarly, α could in principle have up to 4 components if the leakage were anisotropic and involved cross terms[5], but we do not see evidence for anything beyond a real-valued α in practice.

The leakage term $\alpha e_{\rm PSF}$ is commonly (e.g. Heymans et al. 2006) implicitly folded into the general additive error term, c, but we have found it useful to retain it explicitly, since PSF leakage can be one of the more difficult additive errors to correct. Furthermore, Mandelbaum et al. (2015) found that the additive systematic errors for essentially all of the methods submitted to the GREAT3 challenge were well-described by $\alpha e_{\rm PSF}$, which motivates us to include it as an explicit term in equation 3-4.

3.1. Shear Correlation Functions

We set our requirements on the various kinds of systematic errors according to how they propagate into the shear two-point correlation functions (defined as in Jarvis et al. 2003):

$$\xi_+(\theta) = \langle e^*(\mathbf{x})e(\mathbf{x} + \boldsymbol{\theta})\rangle \quad (3\text{-}5)$$

$$\xi_-(\theta) = \langle e(\mathbf{x})e(\mathbf{x} + \boldsymbol{\theta}) \exp(-4i \arg(\boldsymbol{\theta}))\rangle\,, \quad (3\text{-}6)$$

where * indicates complex conjugation.

Substituting equation 3-4 into these equations and assuming the three types of systematic errors are uncorrelated (which is not necessarily true in general, but is a reasonable assumption for setting requirements), we find

$$\delta\xi_i(\theta) \simeq 2m\xi_i(\theta) + \alpha^2\xi_i^{pp}(\theta) + \xi_i^{cc}(\theta) \quad (3\text{-}7)$$

to leading order in each type of systematic, where $i \in \{+, -\}$, $\delta\xi_i$ are the systematic errors in the two correlation functions, ξ_i^{pp} are the auto-correlation functions of the PSF shapes, and ξ_i^{cc} are the auto-correlation functions of the additive error, c.

To set requirements on $\delta\xi_i$, we consider how the errors will affect our estimate of the cosmological parameter σ_8, the present day amplitude of the (linear) matter power spectrum on the scale of $8h^{-1}$ Mpc. Our requirement is that the systematic errors change the estimated value of σ_8 by less than 3%, $\delta\sigma_8/\sigma_8 < 0.03$. This value was chosen to be about half of the expected statistical uncertainty on σ_8 for the DES SV survey.

Propagating this limit to the shear correlation functions, we obtain the requirement

$$\delta\xi_i^{\rm max} = \frac{\partial\xi_i}{\partial\sigma_8}\delta\sigma_8. \quad (3\text{-}8)$$

This constraint assumes that errors are fully correlated across θ; assuming independent errors would be less restrictive.

Figure 3 shows the resulting requirements for $\delta\xi_i$ derived for a flat ΛCDM central cosmological model with $\sigma_8 = 0.82$, $\Omega_b = 0.047$ and $\Omega_c = 0.2344$, $h = 0.7$ and $n_s = 0.96$.

If our requirements are not met for the combination of all potential systematic errors, then their impact will need to be carefully propagated and accounted for in the next stage(s) of the analysis.

3.2. Multiplicative and Additive Errors

From equation 3-7, we find that the requirement on the multiplicative bias, m, is

$$|m| < \frac{1}{2}\left|\frac{\delta\xi_i^{\rm max}}{\xi_i}\right|. \quad (3\text{-}9)$$

[5] In the complex formulation we are using, this would involve terms $\alpha e_{\rm PSF} + \alpha' e^*_{\rm PSF}$. In formulations that treat $[e_1, e_2]$ as a vector, α would be a 2×2 matrix.

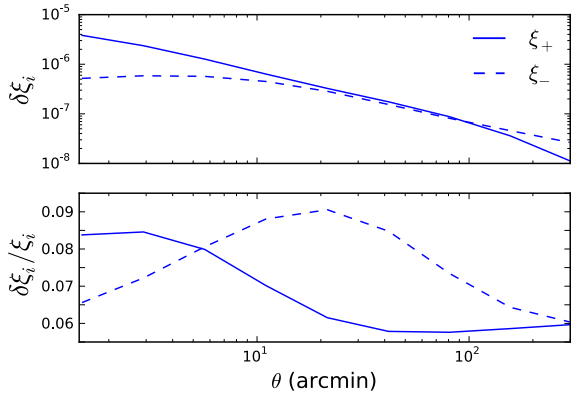

Fig. 3.— Requirement for the maximum systematic error contribution to the shear correlation functions. The blue lines correspond to $\delta\sigma_8/\sigma_8 = 0.03$ for each of the correlation functions ξ_+ (solid) and ξ_- (dashed). The top and bottom panels shows the requirement for the absolute and relative error in the correlation functions.

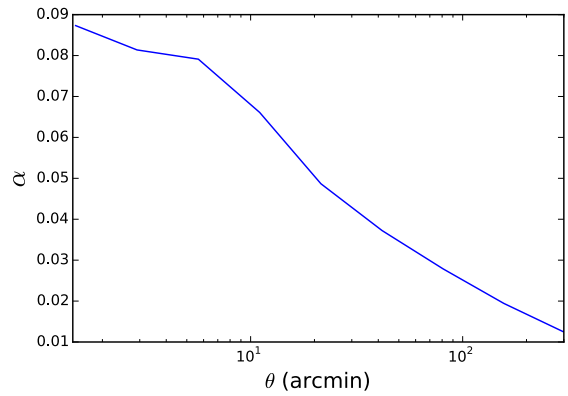

Fig. 4.— Requirement for the PSF leakage factor α based on the relative error in σ_8 being less than 3%.

As can be seen from the lower panel in Figure 3, the most stringent requirement on $\delta\xi_i/\xi_i$ is about 0.06, yielding a requirement on the multiplicative error of

$$|m| < 0.03. \tag{3-10}$$

The requirement on the additive systematic error is somewhat more complicated, since it is the correlation function of the additive systematic that matters. For a systematic error that is coherent over small spatial scales (less than ~ 1 arcminute), the requirement comes from the zero-lag value of $\delta\xi_+$ in Figure 3, $\langle c^2\rangle < \delta\xi_+^{\max}(0)$, or

$$c_{\rm rms} < 2\times 10^{-3}. \tag{3-11}$$

For additive errors that have longer correlation lengths, we will need to be more careful about calculating the correlation function of the systematic error. The requirement in this case is

$$|\xi_i^{cc}(\theta)| < \delta\xi_i^{\max}(\theta) \tag{3-12}$$

using the function shown in Figure 3. The most notable example of this will be systematic effects due to the PSF: both leakage and modeling errors, which will be discussed in the next two sections.

Note that we do not need to satisfy these requirements for all values of θ. The statistical uncertainties on $\xi_{+,-}(\theta)$ become much larger at large scales, so such scales are not as important for constraining cosmology as smaller scales. In practice, equation 3-12 should ideally be satisfied for scales $\theta < 100$ arcminutes, where $\xi_{+,-}(\theta)$ are relatively well-measured.

We note that these results are broadly consistent with those of Amara & Réfrégier (2008), who derived requirements for a tomographic weak lensing survey, performing joint constraints on the set of cosmological parameters for a wCDM model. They found requirements of $|m| < 4.0\times 10^{-2}$ and $c_{\rm rms} < 2.1\times 10^{-3}$ for DES SV survey parameters, which are in rough agreement with the requirements quoted above.

3.3. PSF Leakage

The requirements for the PSF leakage term in equation 3-4 can be obtained from the general requirement on additive errors, equation 3-12.

$$\alpha^2\xi_i^{pp}(\theta) < \delta\xi_i^{\max}(\theta), \tag{3-13}$$

which can be solved for α as

$$|\alpha| < \left(\frac{\delta\xi_i^{\max}(\theta)}{\xi_i^{pp}(\theta)}\right)^{\frac{1}{2}}. \tag{3-14}$$

Figure 4 shows this requirement on α as a function of θ using the observed ξ_+^{pp} for DES SV data[6]. The requirement arising from ξ_-^{pp} is always larger than 0.05 and is not shown.

In general, the amount of leakage of PSF shapes into galaxy shapes from an imperfect correction scheme is not expected to vary with scale. Rather, we can use Figure 4 to determine a conservative requirement for α

[6] See Figure 19, top panels. We use the IM3SHAPE measurement of ξ_+^{pp} here.

that would be applicable for scales $\theta < 100$ arcminutes:

$$|\alpha| < 0.03. \tag{3-15}$$

We will estimate α from the data in §8.2.

3.4. PSF Model Errors

We now consider errors in the modeling of the PSF itself. The previous section dealt with the possibility of the galaxy shear estimation algorithm imperfectly accounting for the PSF convolution and letting some of the PSF shape leak into the galaxy shape. However, even a perfect PSF correction scheme can suffer systematic biases if the PSF model itself is biased.

As our starting point, we use the unweighted moments approximation of Paulin-Henriksson et al. (2008), who give the bias on the measured galaxy ellipticity in terms of errors in the PSF model (their equation 13)[7]:

$$\delta e_{\rm sys} = (e - e_{\rm PSF}) \left(\frac{T_{\rm PSF}}{T_{\rm gal}}\right) \frac{\delta T_{\rm PSF}}{T_{\rm PSF}} - \left(\frac{T_{\rm PSF}}{T_{\rm gal}}\right) \delta e_{\rm PSF}, \tag{3-16}$$

where $T \equiv I_{xx} + I_{yy}$ is the intensity-weighted second moment of the radius (written as R^2 in their paper). $T_{\rm gal}$ refers to the intrinsic galaxy size, unconvolved by the PSF.

Constructing the shear correlation function with this model, we find that the systematic error in ξ_+ is

$$\begin{aligned}\delta\xi_+(\theta) = 2\left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \frac{\delta T_{\rm PSF}}{T_{\rm PSF}} \right\rangle \xi_+(\theta) + \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^2 \rho_1(\theta) \\ - \alpha \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle \rho_2(\theta) + \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^2 \rho_3(\theta) \\ + \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^2 \rho_4(\theta) - \alpha \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle \rho_5(\theta),\end{aligned} \tag{3-17}$$

where $\rho_1(\theta)$ and $\rho_2(\theta)$ are defined as (cf. Rowe 2010)

$$\rho_1(\theta) \equiv \langle \delta e^*_{\rm PSF}(\mathbf{x}) \delta e_{\rm PSF}(\mathbf{x} + \boldsymbol{\theta}) \rangle \tag{3-18}$$

$$\rho_2(\theta) \equiv \langle e^*_{\rm PSF}(\mathbf{x}) \delta e_{\rm PSF}(\mathbf{x} + \boldsymbol{\theta}) \rangle, \tag{3-19}$$

and we introduce three new statistics defined as[8]

$$\rho_3(\theta) \equiv \left\langle \left(e^*_{\rm PSF} \frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\mathbf{x}) \left(e_{\rm PSF} \frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\mathbf{x} + \boldsymbol{\theta}) \right\rangle \tag{3-20}$$

$$\rho_4(\theta) \equiv \left\langle \delta e^*_{\rm PSF}(\mathbf{x}) \left(e_{\rm PSF} \frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\mathbf{x} + \boldsymbol{\theta}) \right\rangle \tag{3-21}$$

$$\rho_5(\theta) \equiv \left\langle e^*_{\rm PSF}(\mathbf{x}) \left(e_{\rm PSF} \frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\mathbf{x} + \boldsymbol{\theta}) \right\rangle. \tag{3-22}$$

There are corresponding terms for $\delta\xi_-$, which are negligible in practice and thus uninteresting as requirements.

The first term in equation 3-17 is a multiplicative systematic, so the relevant requirement comes from equation 3-10. We approximate the ensemble average as a product of two averages to set a requirement on the mean error in the PSF size

$$\left| \left\langle \frac{\delta T_{\rm PSF}}{T_{\rm PSF}} \right\rangle \right| < 0.03 \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^{-1}. \tag{3-23}$$

This represents an error due to improperly accounting for the "dilution", the amount by which the blurring of the PSF makes objects rounder than they originally were. Estimating the wrong PSF size will lead to a systematic multiplicative bias in the inferred galaxy shapes.

The other terms are additive errors, contributing to $\xi_+^{cc}(\theta)$, so the requirements from equation 3-12 are that each term be less than $\delta\xi_+^{\rm max}(\theta)$:

$$|\rho_{1,3,4}(\theta)| < \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^{-2} \delta\xi_+^{\rm max}(\theta) \tag{3-24}$$

$$|\rho_{2,5}(\theta)| < |\alpha|^{-1} \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^{-1} \delta\xi_+^{\rm max}(\theta). \tag{3-25}$$

We will test these requirements for our PSF model below in §4.4.

For our data, we compute the factor $\langle T_{\rm PSF}/T_{\rm gal} \rangle$ that appears in these requirements to be 1.20 for IM3SHAPE and 2.42 for NGMIX; the latter is larger because the final galaxy selection for the NGMIX catalog keeps more small galaxies than the IM3SHAPE selection. We use

[7] The Paulin-Henriksson et al. (2008) formalism is based on $\epsilon = (a^2 - b^2)/(a^2 + b^2)$ rather than our e shape measure, so there are factors of $O(1)$ differences that we are neglecting. Similarly, they derive their formula for unweighted moments, which are also not directly applicable to real shear estimation algorithms, differing again by factors of $O(1)$. Despite these possible shortcomings, we feel this is nonetheless a useful model for describing PSF modeling errors.

[8] We note that Melchior et al. (2015) proposed a slightly different ρ_3 statistic,

$$\rho_3'(\theta) = \left\langle \left(\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\mathbf{x}) \left(\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\mathbf{x} + \boldsymbol{\theta}) \right\rangle,$$

pulling the $e_{\rm PSF}$ factors out of the ensemble average. We believe it is more appropriate to leave them in, since errors in the size estimates could easily be coupled to the PSF shapes.

the NGMIX value in §4.4, as it gives the more stringent requirement. For α, we conservatively use the value 0.03. We will find in §8.2.2 that both codes estimate α to be consistent with zero; however, it is not estimated much more precisely than this value.

4. PSF Estimation

The principal confounding factor that needs to be addressed for measuring accurate shears is the convolution of the galaxy surface brightness profiles by the point-spread function (PSF). The net PSF is due to quite a number of physical processes including atmospheric turbulence, telescope and camera aberrations, guiding errors, vibrations of the telescope structure, and charge diffusion in the CCDs, among other more subtle effects. Furthermore, this PSF is not constant, either spatially over the focal plane, or temporally from one exposure to the next. The atmospheric component varies approximately according to a Kolmogorov turbulent spectrum. The optical aberrations have characteristic patterns that come from standard optical physics.

Fortunately, we do not need to have a complete physical model of all the contributors to the PSF in order to accurately characterize it. Instead, we build an empirical model, based on observations of stars, which we interpolate to obtain an estimate of the PSF at any location on the focal plane. In this section, we describe how we select appropriate stars to use, and then build and test the PSF model.

4.1. Initial Identification of Stars

We first select the stars to use for constraining the PSF model. As stars are point sources, observations of them directly give us an image of the PSF at the location of each star. We want a high-purity sample of fairly bright stars to make sure we do not erroneously consider images of small, faint galaxies to be images of the PSF, as that would bias the resulting PSF model.

We found that the sets of stars identified by the Modest Classification scheme[9] for each CCD image either wrongly include too many galaxies or do not contain enough stars or both. We developed our own algorithm that is more tailored to this purpose. Our algorithm works on each CCD image separately, using a size-magnitude diagram of all the objects detected on the image. For the magnitude, we use the SEXTRACTOR measurement `MAG_AUTO`. For the size, we use the scale size, σ, of the best fit elliptical Gaussian profile using an adaptive moments algorithm. We found that this produces a stellar locus that is flatter and tighter than the `FLUX_RADIUS` value output by SEXTRACTOR, so it works better for this particular purpose. We also initialize the algorithm with some stars identified by SEXTRACTOR to have `CLASS_STAR` between 0.9 and 1.0. This was found to give a decent estimate of the size of the PSF, providing a good starting guess for the location of the stellar locus.

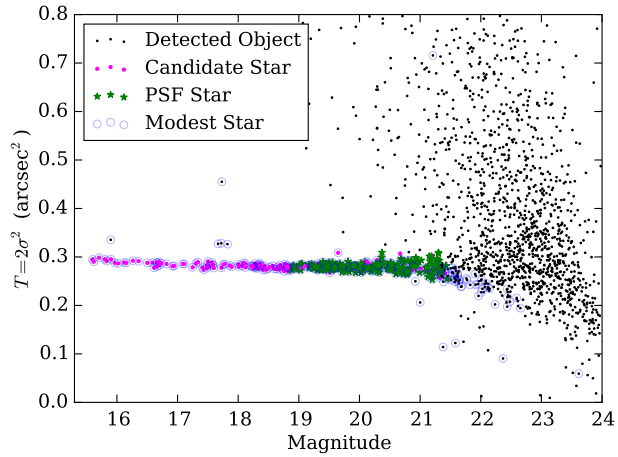

Fig. 5.— An example size-magnitude diagram for a single CCD image, used to identify stars. The size $T = 2\sigma^2$ is based on the scale size of the best fit elliptical Gaussian. The pink and green points are the objects initially identified as stars. The green points are the ones that pass our selection criteria outlined in §4.2, most notably the magnitude cut to avoid objects contaminated by the brighter-fatter effect. These objects are then used to constrain the PSF model. The blue circles show an an alternate star classification, called the Modest Classification within DES, which was found not to work as well for our specific purpose.

The stars are easily identified at bright magnitudes as a locus of points with constant size independent of magnitude. The galaxies have a range of sizes, all larger than the PSF size. Thus, the algorithm starts with a tight locus at small size for the stars and a broad locus of larger sizes for the galaxies for objects in the brightest 5 magnitudes (excluding saturated objects). Then the algorithm proceeds to fainter magnitudes, building up both loci, until the stellar locus and the galaxy locus start to merge. The precise magnitude at which this happens is a function of the seeing as well

[9] Stars are identified as `(bright_test OR locus_test)` in terms of the pseudo-code presented in §2.2

as the density of stars and galaxies in the particular part of the sky being observed. As such the faint-end magnitude of the resulting stellar sample varies among the different exposures.

Figure 5 shows such a size-magnitude diagram for a representative CCD image. The stellar locus is easily identified by eye, and the stellar sample identified by our algorithm is marked in pink and green. The pink points are stars that are removed by subsequent steps in the process outlined below, while the green points are the stars that survive these cuts. The blue circles show the objects identified as stars according to the Modest Classification, which includes more outliers and misses some of the objects clearly within the stellar locus.

While the algorithm we currently use is found to work well enough for the SV data, we plan to investigate whether the neural net star-galaxy separator recently developed by Soumagnac et al. (2015) is more robust or could let us include additional stars.

4.2. Selection of PSF Stars

Some of the stars in this sample are not appropriate to use to model the PSF, even ignoring the inevitable few galaxies that get misidentified as stars. The CCDs on the Dark Energy Camera each have six spots where 100 micron thick spacers were placed behind the CCDs when they were glued to their carriers (cf. Flaugher et al. 2015), which affects the electric field lines near each 2mm $\times$ 2mm spacer. These features, which we call tape bumps, distort the shapes in those parts of the CCDs, so the stellar images there are not accurate samples of the PSF. We exclude any star whose position is within 2 FWHM of the outline of each tape bump. The tape bumps are small, so this procedure excludes less than 0.1% of the total area of the CCD, but removes a noticable bias in the PSF model near the bumps.

The other problem we need to address with the star selection is known as the "brighter-fatter effect" (Antilogus et al. 2014; Guyonnet et al. 2015). As charge builds up in each pixel during the exposure, the resulting lateral electric fields and increased lateral diffusion push newly incoming charges slightly away from the existing charge. This makes bright objects a bit larger than fainter objects. In addition, an asymmetry in the magnitude of the effect between rows and columns can make bright stars more elliptical. The galaxies we will use for weak lensing are generally faint, so the brightest stars do not accurately sample the PSF that we need to measure. Furthermore, the brighter-fatter effect does not manifest as a convolution of the signal, so the bright stars do not even provide an estimate of the correct PSF to be used for bright galaxies.

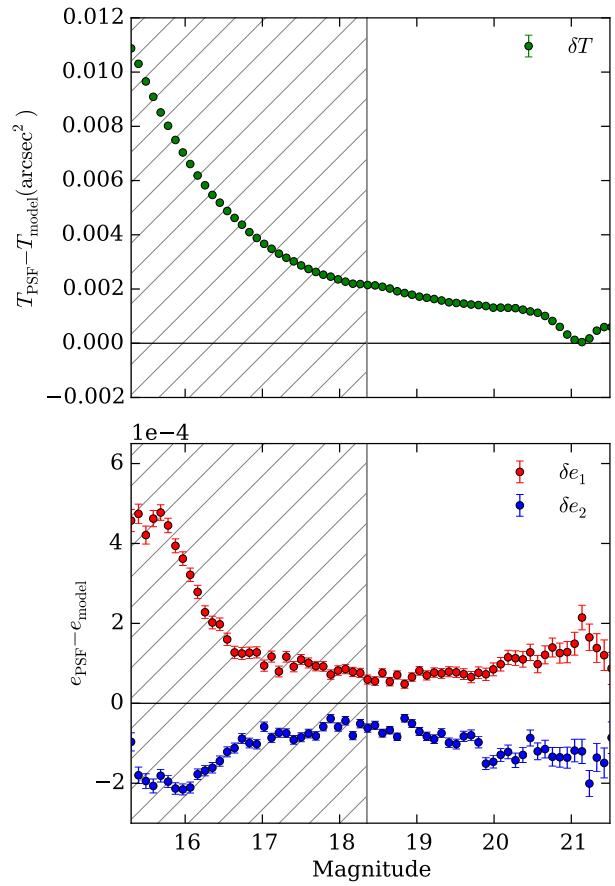

Fig. 6.— The residual size (top) and shape (bottom) of stars relative to that of the PSF model as a function of magnitude. The hatched region on the left shows the magnitude range of the stars we exclude from the sample to reduce the impact of the brighter-fatter effect.

The appropriate solution is to move the shifted charge back to where it would have fallen in the absence of this effect. This will be implemented in future DES data releases (Gruen et al. 2015). For the current round of catalogs, we instead try to avoid the problem by removing the brightest stars from our sample. Specifically, we removed all stars within 3 magnitudes of the saturation limit for the exposure. That is, the brightest pixel in the stellar image must be less than 6% of the pixel full well for us to use the star as a sample of the PSF. Since the brighter-fatter effect scales linearly with flux, this reduces the magnitude of the effect by a factor of 16. We are left with stars of lower S/N, so it is not the ideal solution, but it is an acceptable interim measure (as we demonstrate below) until

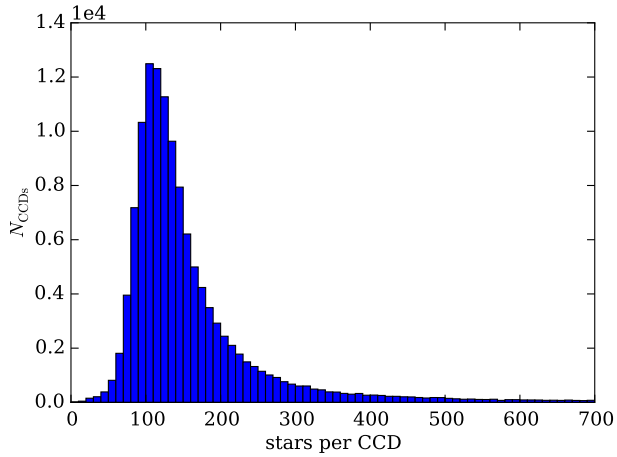

Fig. 7.— The distribution of the number of stars per CCD image used for constraining the PSF model.

the more sophisticated solution can be implemented.

Figure 6 shows the mean difference between the measured sizes of observed stars and the size of the PSF model at their locations, using the model described below in §4.3. For the measurements of the sizes and shapes described here, we use the implementation of the HSM (Hirata & Seljak 2003; Mandelbaum et al. 2005) algorithm in GALSIM. The hatched region shows the range we exclude to avoid the spurious increase in PSF size from the brighter-fatter effect. Figure 6 also shows the mean difference in ellipticity due to the brighter-fatter effect; it affects the shapes of the stars as well as the size.

It is not understood why the residual sizes and shapes shown in Figure 6 do not level off at zero at fainter magnitudes where the brighter-fatter effect is negligible. The requirement on this residual value is given by equation 3-23. We calculate $\langle \delta T_{\rm PSF}/T_{\rm PSF} \rangle$ to be 0.0044, which is well below our requirement of 0.013 for the SV data. However, this residual will not be acceptable for future DES analyses, so we will need to investigate what is causing the problem and fix it.

The complete process described above finds a median of 130 stars per CCD image to use for constraining the model of the PSF. The distribution is shown in Figure 7.

Sometimes a CCD image will end up with no selected stars, either because the initial stellar selection could not find any stars or no stars survived the magnitude cuts. For instance, this can happen when there is a very bright object in the image that essentially masks out the entire image, leaving no or very few objects detected. In less extreme cases, the bright object can contaminate the fluxes and sizes of the other detections sufficiently that the stellar locus is either difficult to find or merges with the galaxy locus at a fairly bright magnitude so the brighter-fatter cut excludes the entire sample.

Whenever the process fails for any reason, we flag the CCD image that failed and exclude it from being used in subsequent shear estimation. We also flag images with less than 20 identified PSF stars, since it is difficult to accurately interpolate the PSF model with so few stars. These flagged images are added to the set of blacklisted images described in §5.1.

4.3. PSF Measurement and Interpolation

To measure the PSF and its spatial variation on each CCD, we use the software package PSFEX (Bertin 2011). Normally, PSFEX takes as input the full list of objects detected by SEXTRACTOR and finds the bright stars automatically. However, as described in §4.2, we remove some of the stars in the catalog to avoid the brighter-fatter effect and the tape bumps. This edited catalog of stars is then passed to PSFEX.

We use the `BASIS_TYPE = PIXEL_AUTO` option, which uses pixelated images to model the PSF profile, rather than fitting to some functional form. In Kitching et al. (2013), for undersampled PSFs a fixed oversampling was found to perform better than the default PSFEX choice; therefore, we force the oversampling of these images to be a factor of 2 finer than the original pixel size with `PSF_SAMPLING = 0.5`. The basis images are set to be 101×101 in the resampled pixels, or approximately 13 arcseconds on a side.

For the interpolation, we use second order polynomials in chip coordinates, interpolated separately on each CCD. Specifically, we use the following parameters:

```
PSFVAR_KEYS = XWIN_IMAGE,YWIN_IMAGE
PSFVAR_GROUPS = 1,1
PSFVAR_DEGREES = 2
```

We found that there was not much gain in using higher order polynomials than this and some evidence that they were overfitting the noise for some CCDs. So we decided to use second order in all cases.

To assess the quality of the PSF interpolation, we first look at the differences between the measured shapes (using the HSM algorithm again) of actual stars on the image and the corresponding values for

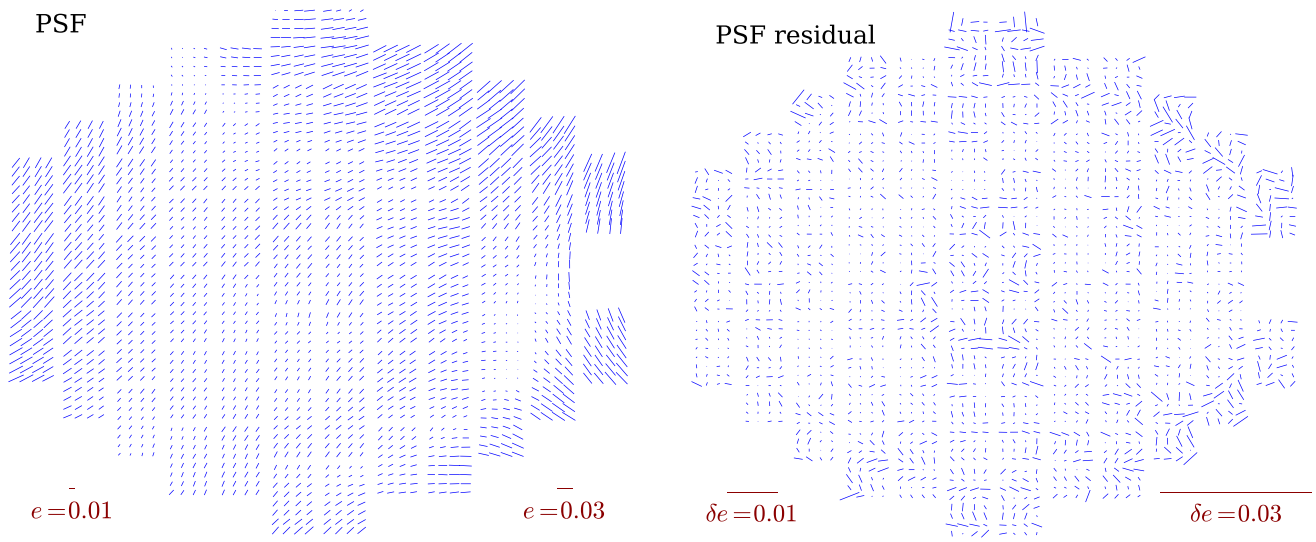

Fig. 8.— Whisker plots of the mean PSF pattern (left) and of the mean residual after subtracting off the model PSF (right) as a function of position in the focal plane. The length of each whisker is proportional to the measured ellipticity, and the orientation is aligned with the direction of the ellipticity. There is still some apparent structure in the plot of the residuals, but the level is below the requirements for SV science. Reference whiskers of 1% and 3% are shown at the bottom of each plot, and we have exaggerated the scale on the right plot by a factor of 10 to make the residual structure more apparent.

the PSFEx model at the locations of these stars. Figure 8 shows whisker plots of the PSF and the residuals as a function of position on the focal plane. The residual whiskers are small, but not quite zero. The impact of these spatially correlated residuals are investigated below in §4.4, and we will see that they meet the requirements for science with SV data.

We believe the remaining structure seen in the residual plot is largely due to the fact that the PSF modeling and interpolation is done in pixel coordinates rather than sky coordinates. Therefore, the interpolation must also include the effects of the non-uniform WCS. In particular, the distortion due to the telescope optics is a fifth order radial function, but we fit the PSF with only a second order polynomial on each CCD. This is most markedly seen in the CCDs near the edges of the field of view where the residuals look consistent with a fifth order radial function after the local second order approximation has been subtracted off. One of our planned improvements to the analysis is to change the PSF interpolation to be in sky coordinates, so it does not have to accurately model the WCS variations along with the real PSF variations. We expect this change to remove most of the remaining residual PSF pattern.

4.4. PSF Model Diagnostics

Errors in the PSF model, and particularly errors in the interpolation, will directly affect the shear estimates of galaxies, since they would be accounting for the effect of the PSF convolution incorrectly, as discussed in §3.4. If the PSF errors were random, independent values for each galaxy, they would constitute merely an additional contribution to the shear measurement uncertainty, which would be highly subdominant to other sources of statistical noise, such as the unknown intrinsic shapes of the galaxies. However, this is not the case. Because the PSF is interpolated between stars, the errors in the PSF estimate are correlated among nearby galaxies. The two-point correlation function of these errors will directly impact the two-point correlation function of the shear estimates, which means they would be a systematic error, as quantified in equation 3-17.

Rowe (2010) describes two diagnostic functions to quantify the level of interpolation errors in the PSF model using the measured shapes of stars and the interpolated value of the model at the locations of these stars. As we already introduced in §3.4,

$$\rho_1(\theta) \equiv \langle \delta e^*_{\rm PSF}(\mathbf{x}) \delta e_{\rm PSF}(\mathbf{x} + \boldsymbol{\theta}) \rangle \qquad (4\text{-}1)$$

$$\rho_2(\theta) \equiv \langle e^*_{\rm PSF}(\mathbf{x}) \delta e_{\rm PSF}(\mathbf{x} + \boldsymbol{\theta}) \rangle \,, \qquad (4\text{-}2)$$

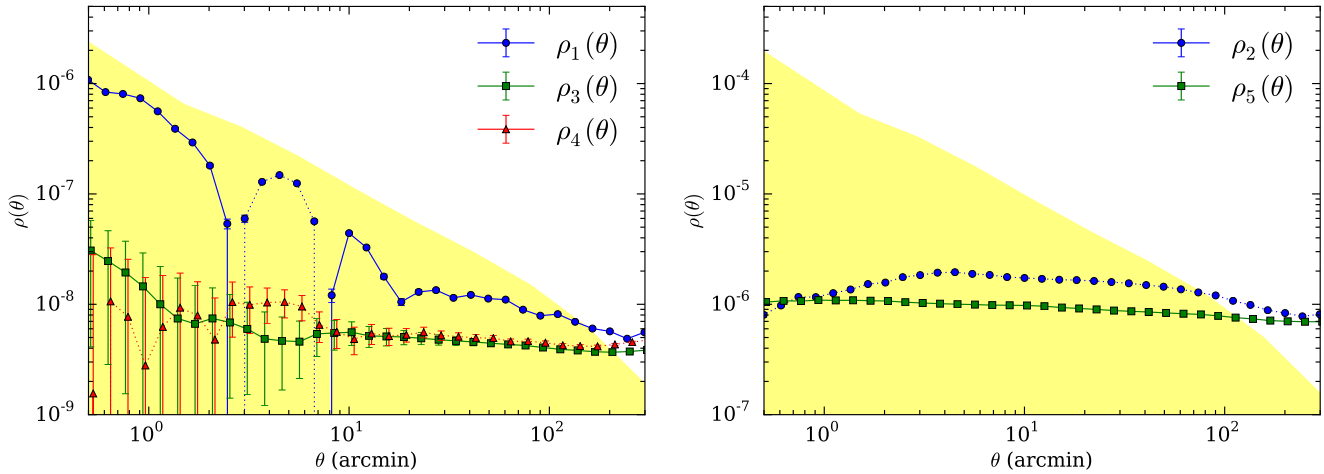

Fig. 9.— The ρ statistics for the PSF shape residuals. Negative values are shown in absolute value as dotted lines. The shaded regions are the requirements for SV data.

where $\delta e_{\rm PSF}$ represents the difference between the measured ellipticity of the observed stars and the ellipticity of the PSFEx models at the same locations, which is an estimate of the systematic uncertainty in the shape of the PSF model at those locations.

In addition, we test three other statistics that appear at the same order of the expansion of PSF model errors, involving errors in the PSF size, $T_{\rm PSF}$, which we call ρ_3, ρ_4, and ρ_5. They are defined in equations 3-20 – 3-22 and are generally smaller than the two described by Rowe (2010).

Figure 9 shows the results for these five statistics. The shaded regions show the requirements, from equations 3-24 & 3-25. In all cases, the results are seen to be passing our requirements for scales less than about 100 arcminutes. We see in Figure 9 that ρ_1 changes sign twice and is below the requirement line by only a factor of ~ 2. However, our requirements make the conservative assumption that additive errors are fully correlated across scales. So we have directly propagated the measured ρ_1 through to the bias on σ_8 and found the influence on σ_8 to be much less than one percent.

Of course, the PSFEx model describes the full surface brightness profile of the PSF, not just its shape $e_{\rm PSF}$ and size $T_{\rm PSF}$. However, these are the dominant ways that errors in the PSF model could affect the galaxy shapes, so these statistics are the most important checks of the model accuracy.

5. Multi-Epoch Data Structures

As outlined in §2, we use the coadd images for object detection and deblending. For shear measurement we work directly with the pixel data from the original single-epoch images (cf. §7.1). To simplify the bookkeeping we developed a new data storage format, which we named Multi-Epoch Data Structures (MEDS)[10]

We make a MEDS file corresponding to each coadd image. This file stores a postage stamp for each observation of every object detected in the coadd image along with the corresponding weight maps, segmentation maps, and other relevant data. The postage stamps for each coadd object are stored contiguously in the file, making sequential access of individual objects efficient. The files are quite large, so loading the whole file into memory is not generally feasible, but it is also not necessary.

The postage stamps from the original single-epoch images are sky-subtracted and then scaled to be on a common photometric system, which simplifies the model fitting using these images. We also store the local affine approximation of the WCS function, evaluated at the object center, so that models can be made in sky coordinates and constrained using the different image coordinates for each postage stamp.

See Appendix A for full details about how we build and store the MEDS files.

[10] https://github.com/esheldon/meds

5.1. Exposure Selection

We do not use all single-epoch images for measuring shapes. We exclude a small fraction of the CCD images that have known problems from some part of the data reduction process. We create simple "blacklist" files, in which we store information for CCD images that should be excluded, and that information is incorporated into the MEDS files as a set of bitmask flags. Postage stamps from blacklisted images are then easily excluded from the analysis when measuring shears. Here we list some of the reasons that images are blacklisted.

Some of the astrometry solutions (cf. §2.3) fail to produce a good map from CCD coordinates to sky coordinates. This primarily happens near the edges of the SPT-E region where there are not enough overlapping exposures to constrain the fit.

Some of the PSF solutions (cf. §4) fail to produce a good model of the PSF across the CCD. This may be because there were too few stars detected or occasionally because there was some error in either the star finding code or in PSFEx.

A small fraction of the SV images are contaminated by bright scattered-light artifacts. Scattered-light artifacts fall into two broad categories: internal reflections between the CCDs and other elements of the optics, known as "ghosts"; and grazing incidence reflections off of the walls and edges of the shutter and filter changer mechanism. Ghosts primarily occur when a bright star is within the field of view, while grazing incidence scatters occur predominantly for stars just outside the field-of-view. Using the positions of bright stars from the Yale Bright Star Catalog (Hoffleit & Jaschek 1991) and knowledge of the telescope optics, it is possible to predict locations on the focal plane that will be most affected by scattered light. We identified and removed a total of 862 CCD images (out of 135,481) from the single-exposure SV data set in this manner. In April 2013, filter baffles were installed to block some of this scattered light, and non-reflective paint was applied to the filter changer and shutter in March 2014 (Flaugher et al. 2015). These modifications have greatly reduced the occurrence of grazing incidence reflections in subsequent DES seasons.

It is common for human-made objects to cross the large field of view during an exposure. The brightest and most impactful of these are low-flying airplanes (two Chilean flight paths pass through the sky viewable by the Blanco telescope). Airplane trails are both bright and broad, and cause significant issues in estimating the sky background in CCDs that they cross. Airplane trails were identified by eye and affected images were removed by hand. In total, 56 individual CCD images were removed due to airplane contamination (corresponding to 4 distinct exposures). This rate of airplane contamination is expected to continue throughout the DES survey.

In addition to airplanes, earth-orbiting satellites are a common occurrence in DES images. During the 90 second exposure time of a DES survey image, a satellite in low-earth-orbit can traverse the entire focal plane, while geosynchronous satellites travel approximately 1.25 CCD lengths. The impact of these satellite streaks is significantly less than that of airplanes; however, because they only occur in a single filter, they can introduce a strong bias in the color of objects that they cross. For SV, the "crazy colors" cut mentioned in §2.1 removes most of the contaminated objects. At the end of Year 1, an automated tool was developed by DESDM for detecting and masking satellite streaks using the Hough transform (Hough 1959; Duda & Hart 1972). This should greatly reduce the impact of satellite streaks in upcoming seasons of DES observing and will be retroactively applied to reprocessing of earlier data.

5.2. Masks

The user can construct a "mask" for each postage stamp in the MEDS files in a variety of ways. We currently use what we call an "überseg" mask, constructed from the weight maps, segmentation maps and locations of nearby objects.

To create the überseg mask, we start with the SExtractor segmentation map from the coadd image and map it onto the corresponding pixels of the single-epoch images. We do this rather than use the segmentation map derived for each single-epoch image because the coadd image is less noisy and thus has more object detections and more information for determining the extent of each object.

We then set pixels in the weight map to zero if they are either associated with other objects in the segmentation map or are closer to any other object than to the object of interest. What results is a superset of the information found in the weight maps and segmentation maps alone, hence the name überseg.

An example set of images and überseg maps are shown in Figure 10. Tests on a simulation with realistically blended galaxies (cf. §6.2) showed a large im-

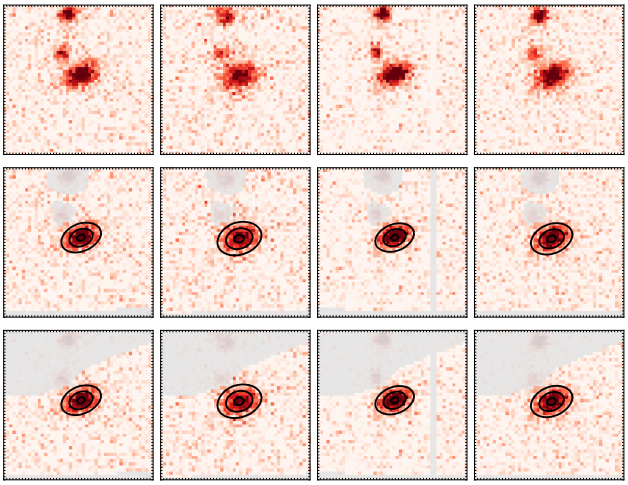

Fig. 10.— Example galaxy image demonstrating two masking strategies. The top row shows the original postage stamps in the MEDS file. The second row shows the result when only the SEXTRACTOR segmentation map was used to mask neighbors. The third row shows the result when the überseg algorithm was used to mask neighbors, as described in the text.

provement in the shear biases from using the überseg masking over just using the SEXTRACTOR segmentation maps. Light from regions that were not masked by the segmentation maps alone caused a significant bias in the shapes of some galaxies toward the direction of a neighbor. With the überseg masking, such a bias was undetectable.

6. Simulations

Simulations are a crucial part of any shear pipeline development process, because they can provide data with known values of the applied shear for testing the shear estimation code. In addition to many small targeted simulations designed to answer particular questions about the algorithms, we developed two more general purpose simulations that we used extensively to test the shear pipelines.

The first, which we call GREAT-DES, is modeled on the GREAT3 challenge. It tests the accuracy of the shear estimation codes on realistic space-based galaxy images with a realistic range of noise levels and galaxy sizes. As with the GREAT3 challenge, the galaxies are placed on postage stamps, so there are no blending or object detection issues to worry about.

The second, which we call the end-to-end simulation, is a high S/N simulation with analytic galaxy models with elliptical isophotes. The point here is to test that various bookkeeping details are implemented correctly, such as the file conventions used by PSFEX, the application of the WCS transformations, and conventions about the origin of the postage stamps in the MEDS files. These are all details that are easy to get wrong, but which can be difficult to notice on noisy data.

We have found the GALSIM (Rowe et al. 2015) image simulation software to be invaluable for this purpose. In particular, its ability to accurately render sheared versions of space-based images using their reimplementation of the SHERA algorithm (SHEar Reconvolution Analysis; Mandelbaum et al. 2012), correctly accounting for the original HST PSF (Bernstein & Gruen 2014), was particularly important for making the GREAT-DES simulation. The end-to-end simulation relied on GALSIM's ability to generate multiple epochs of the same scene and accurately handle non-trivial WCS transformations for the various exposures.

6.1. GREAT-DES

The GREAT-DES simulation tests the precision of shear measurement codes with DES-specific sampling of both the population of galaxies (size, shape, morphology) and the observing conditions (PSF ellipticity, noise level). The simulation consists of individual 48×48 pixel postage stamp images. We ignore issues of crowding, bad pixels, and imaging artifacts, but we otherwise attempt to make the images a close approximation to the DES SV data.

We build the GREAT-DES simulation using galaxies from the COSMOS survey (Koekemoer et al. 2007), made available for use with GALSIM[11]. Kannawadi et al. (2015) showed that this sample of galaxies is a good representation of galaxy properties, and can be used in shear calibration of lensing surveys to a precision level of $m = 0.01$.

We start with the entire COSMOS sample distributed for use with GALSIM and discard objects that were flagged as unusable in the GREAT3 challenge (Mandelbaum et al. 2014), which removes about 3% of the objects and leaves more than 54,000 COSMOS galaxies available to use for the simulation. Next we select individual galaxies from this set according to how common each galaxy's properties are in the DES SV data.

[11] https://github.com/GalSim-developers/GalSim/wiki/RealGalaxy%20Data

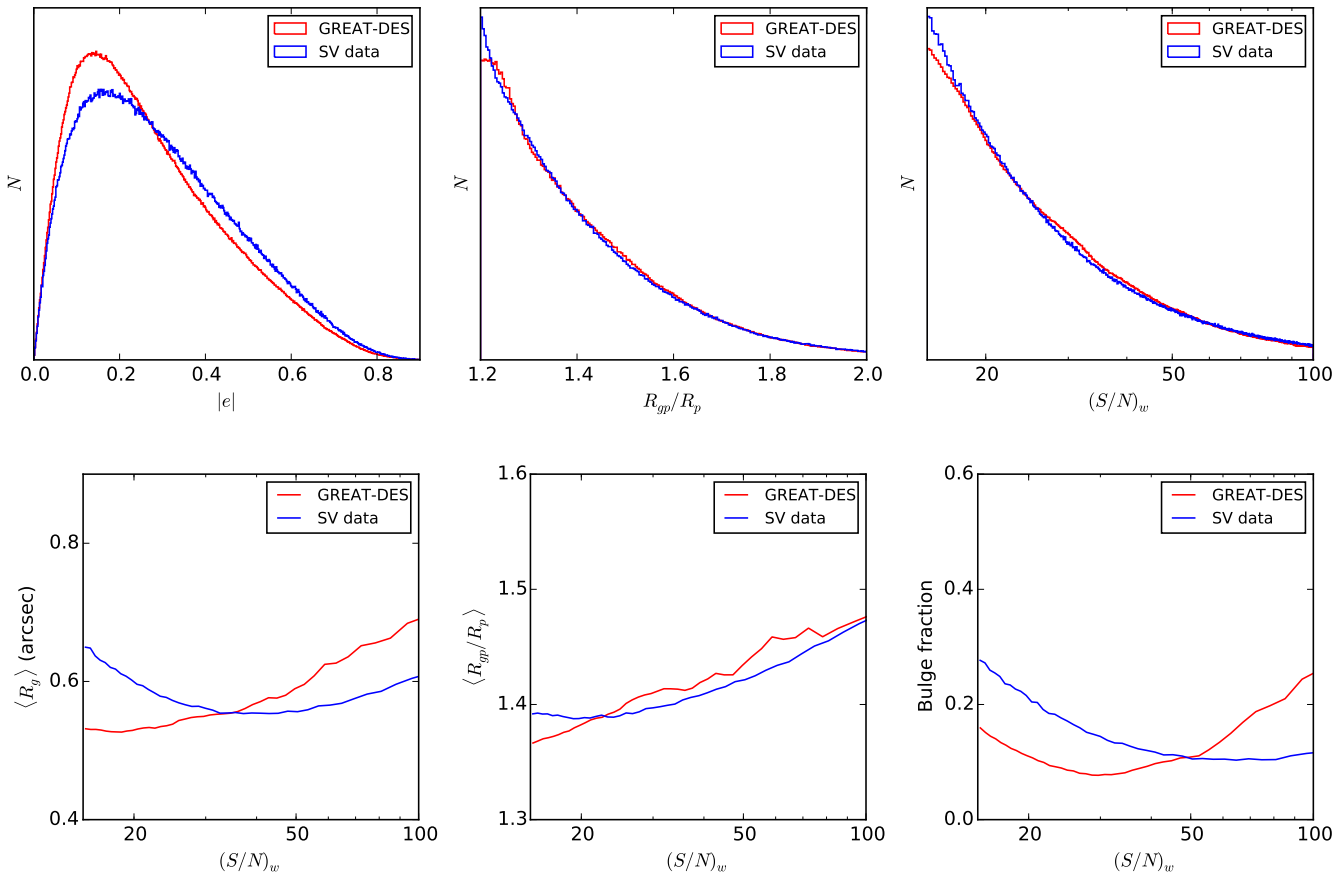

Fig. 11.— A comparison of the galaxy properties in GREAT-DES (red) and the SV data (blue). The top row shows histograms of $|e|$, R_{gp}/R_p, and S/N as measured by IM3SHAPE. The bottom row shows the dependence of $\langle R_g \rangle$, $\langle R_{gp}/R_p \rangle$, and bulge fraction as functions of S/N.

.

For the PSF, we use a Kolmogorov profile with sizes and ellipticities taken to match the range of values present in the SV data. Specifically the PSF size takes one of 6 values between 0.8 and 1.3 arcseconds FWHM, and each component e_1, e_2 of the shape takes one of 4 values from -0.02 to +0.02. Thus, a total of 96 unique PSF images are used in the entire sample. Gaussian noise is added based on the typical noise level observed in SV coadd images.

We then apply constant shear values within each simulation field, with a magnitude of $|g| = 0.05$ and rotated at 8 evenly spaced position angles ϕ.

Each of the COSMOS galaxies is used hundreds of times, with different noise realizations, different random orientations and different centroid offsets. We do not use 90 degrees rotated galaxy pairs, as has commonly been done for the GREAT challenges (cf. Kitching et al. 2010) to reduce the number of galaxies required to reach a given measurement precision, because we need to be able to select subsets of the galaxies according to their measured characteristics. This tends to break up the pairs, which obviates any advantage from using them. Therefore, we instead use a very large number of galaxies in each field to get to the desired precision on the mean shear. In total, we use 48 million rendered galaxy images.

We developed a DES-specific module for GALSIM to store the simulated images directly in MEDS format, so we can run IM3SHAPE and NGMIX on the resulting MEDS files with minimal modification compared to how we run the code on the data. We did not actually create multiple epochs for this simulation, but it was helpful to use the same file format as the data.

To estimate the level of systematic errors well, the simulation must be representative of the data (see e.g. Bergé et al. 2013; Bruderer et al. 2015). To check that we have achieved this goal, we compare the IM3SHAPE measurements of relevant galaxy proper-

ties in GREAT-DES to those in the SV data. Figure 11 shows that the distributions of ellipticity, size and S/N are well-matched between the simulation and the data, as is the dependence of R_g and R_{gp}/R_p as functions of S/N. The bulge fraction (estimated by which model IM3SHAPE chooses as the better fit) does show some disagreement, although the overall bulge fraction matches very well: 0.161 and 0.167 for the simulation and the data respectively.

Note that the choice to show IM3SHAPE measurements in Figure 11 is arbitrary; the analogous plot of NGMIX measurements shows similar agreement, except that there is no estimate of bulge fraction from the NGMIX exponential disk model. Also, the IM3SHAPE "bulge fraction" should not be considered an estimate of the actual Sérsic index of the galaxies; it is merely a diagnostic measure related to the the concentration of the galaxies.

The discrepancy of the bulge fraction as a function of S/N is not expected to significantly impact IM3SHAPE shear estimates, since its bulge-or-disk galaxy model (cf. §7.3.1) can handle galaxies with a range of Sérsic indices with very little model bias (Kacprzak et al. 2014). But since NGMIX uses an exponential disk model (cf. §7.4.1), it has worse model bias for bulge-like galaxies, so the difference may be more relevant for it (cf. §8.5).

6.2. End-to-end Simulation

The end-to-end simulation is of an entirely different nature from the GREAT-DES simulation. It is a high S/N simulation used to test various mundane coding details that are easy to mix up, but which can be difficult to verify in noisy data. For the galaxies we use simple exponential disk profiles, which have elliptical isophotes when sheared, and the images are rendered with very little pixel noise. The fundamental shape estimation problem is thus trivial for both algorithms.

The starting point for this simulation is one of the actual MEDS files from the data, along with the corresponding coadd catalog, the list of single-epoch images that contributed to the coadd image, and the WCS solutions and estimated background maps for each single-epoch image.

Next we build new versions of these single-epoch images using exponential disk galaxies with the same size, flux, ellipticity, and celestial position as the measurements of the real galaxies. We use variable elliptical Gaussian profiles for the PSF, using different parameters for each single-epoch image. The convolved images are rendered at the correct position on each image using the original WCS. GALSIM automatically applies the Jacobian of the WCS to the surface brightness profile as well, so this detail is handled correctly. Objects that were deemed to be stars in the original catalog (based on the SEXTRACTOR `SPREAD_MODEL` being less than 0.003) are drawn as a PSF profile with the same flux as the original object. Finally, we add the original background sky level to the image and add a very small amount of noise so that the faintest galaxies have $S/N > 200$.

We then run these images through the full weak lensing pipeline, starting with SEXTRACTOR and PSFEX to estimate the PSF model, then building a MEDS file, and finally running IM3SHAPE and NGMIX. The resulting measured shape estimates are then compared to the true shapes of the simulated galaxies, which are expected to match to quite high precision, given the nature of the simulation.

The end-to-end simulation was successful in finding several bugs in various parts of the shear pipeline. However, the most notable result from this process was the development of the überseg mask (cf. §5.2). These tests revealed significant biases from the masking procedure we had been using, involving just the SEXTRACTOR segmentation maps. When a galaxy had a bright neighbor on the same postage stamp, light from the neighbor that was just outside the segmentation map was being included as part of the fit, thus significantly biasing the inferred shapes in the direction of the neighbor.

Switching to the überseg mask made a big difference; the measured shapes are now much closer the true values. There is still a small effect from neighbors, which amounts to a slight increase in the effective shape noise for such objects, but we no longer detect any systematic bias in the shape estimates due to unmasked flux from neighboring objects.

7. Shear Measurement

We use two different shear measurement codes for this study: IM3SHAPE and NGMIX, both of which are model-fitting algorithms. IM3SHAPE performs a maximum likelihood fit using a bulge-or-disk galaxy model (cf §7.3.1). NGMIX uses an exponential disk model, exploring the full N-dimensional posterior likelihood surface with an informative prior applied on the ellipticity (cf. §7.4.1).

Both shear methods use the PSFEX models of the

PSF detailed in §4.3, although they use the models differently. PSFEx produces a 2-D image of the PSF profile at the location of each galaxy. IM3SHAPE resamples the PSF image onto a higher resolution grid and performs the convolution with the galaxy model via FFT. NGMIX fits 3 elliptical Gaussians to the PSF image and performs an analytic convolution with the galaxy, which is also modeled as a sum of Gaussians.

Finally, both shear codes use the MEDS files described in §5 for constraining the galaxy models using the original single-epoch pixels rather than the coadd image, which we only use for object detection.

We discuss the details of the multi-epoch fitting process in the next section, §7.1. In §7.2, we define what we mean by signal-to-noise. The details of the IM3SHAPE and NGMIX algorithms are given in §7.3 and §7.4. Finally, our strategy for blinding the shear estimates is described in §7.5.

7.1. Multi-epoch Fitting

The typical method for dealing with multiple exposures of a particular patch of sky is co-addition of images (also known as "stacking"; cf. Fu et al. 2008). However, co-addition can be problematic, since it necessarily loses information and imparts non-trivial, spatially correlated noise into the final image. Furthermore, as each CCD covers a finite region of sky, addition of a finite number of CCD images results in discontinuities in the PSF at image boundaries (cf. Jee et al. 2013 for a discussion of this effect). A more optimal method for fitting a collection of images is to simultaneously fit all independent pixel data, as also advocated by Heymans et al. (2012b). We call this process multi-epoch fitting.

Multi-epoch fitting requires some additional complications to the fitting process, as we must use the correct PSF and WCS information for each image, rather than a single function for each as would be sufficient to process a coadd image.

In order to simplify the bookkeeping to process the multi-epoch and multi-band DES data, we use the Multi-Epoch Data Structures described in §5. Each observation of a particular galaxy experiences a different PSF, and the local image coordinates are related to celestial coordinates via a different WCS transformation. This information is stored in the MEDS file and used during modeling.

For both codes (NGMIX and IM3SHAPE) the model for a given set of galaxy parameters is generated in celestial coordinates. For NGMIX the PSF is modeled in celestial coordinates as well and is convolved with the galaxy model analytically. This is then compared to the observed data using the WCS transformation. For IM3SHAPE the galaxy and PSF models are rendered in image coordinates and convolved via fast Fourier transform (FFT).

7.2. Signal-to-noise Ratio

Before we describe the algorithms we use for measuring shapes, it is worth describing in detail what we mean by the signal-to-noise ratio (S/N). This will be relevant both in the next section (in particular §7.3.2, where we discuss how IM3SHAPE calibrates the shear bias) and in later sections such as §8.3, where we test that the shear is independent of S/N, and §9.1, where we use S/N while selecting galaxies for the final shear catalogs.

There is no single definition for the S/N of an image or a surface brightness profile. Rather, a S/N is only well defined for a single measured value – some statistic calculated from the image or profile. Given some such statistic x, the S/N is typically defined as that value (either the measurement or the true value) divided by the square root of its variance

$$S/N \equiv \frac{x}{\sqrt{\mathrm{Var}(x)}}. \tag{7-1}$$

One of the standard S/N measures is the so-called "optimal" S/N estimator. One can show that among all statistics that are linear in the pixel values I_p,

$$\hat{I}_w = \sum_p w_p I_p, \tag{7-2}$$

the one with the highest expected S/N has weights $w_p = \langle I_p \rangle / \sigma_p^2$, where σ_p^2 are the estimated variances in each pixel.[12]

In practice, one does not know the true expectation value of the surface brightness profile, $\langle I_p \rangle$, so typically one uses the best fit model of the galaxy, which we call m_p, as part of the weight. The S/N of this statistic is thus estimated as

$$(S/N)_w = \frac{\sum_p m_p I_p / \sigma_p^2}{\left(\sum_p m_p^2 / \sigma_p^2\right)^{1/2}}. \tag{7-3}$$

[12] The proof involves finding w_p values such that expectation of the S/N is stationary with respect to any infinitesimal changes δw_p.

This is the S/N measure used by GREAT3 (Mandelbaum et al. 2014), for example.

A drawback of this estimator is that it is not independent of an applied shear. Galaxies that look similar to the PSF will have a higher measured $(S/N)_w$ than galaxies with a different size or shape. The PSF essentially acts as a matched filter for these galaxies. This means that $(S/N)_w$ is not invariant under an applied gravitational shear.

If the PSF is approximately round, as is the case for our data, then more elliptical galaxies will have a lower estimated $(S/N)_w$ than round galaxies (holding flux constant). Thus if galaxies are selected according to their measured $(S/N)_w$, the resulting galaxy catalog will have a selection bias towards round shapes, which will bias the overall mean shear.

One solution to this potential systematic error is to use a S/N estimator that is not biased with respect to an applied shear. There are a number of choices one could make for this. We choose to calculate the $(S/N)_w$ that the galaxy would have had *if it and the PSF were round*.

That is, we take the model of the galaxy profile and apply a shear such that its ellipticity becomes zero. We do the same for the PSF, convolve these two profiles together, and then integrate over the pixels. The resulting m_p^r values are the intensities we predict would have been observed if both the galaxy and the PSF had been round. We then use these values for both the model m_p and the intensity I_p in equation 7-3, as the actual data are no longer appropriate for this counterfactual surface brightness profile. The "roundified" S/N estimator is then

$$(S/N)_r = \frac{\sum_p m_p^r m_p^r/\sigma_p^2}{\left(\sum_p (m_p^r)^2/\sigma_p^2\right)^{1/2}} = \left(\sum_p (m_p^r)^2/\sigma_p^2\right)^{1/2}. \qquad (7\text{-}4)$$

We find both measures of the signal-to-noise useful in different contexts. For NGMIX, we use $(S/N)_r$ for the reasons described here; we find significantly smaller selection biases when we use $(S/N)_r$ to select galaxies for shear measurement, as compared to using $(S/N)_w$.

For IM3SHAPE, we find that the noise bias calibration (cf. §7.3.2) is more accurate using $(S/N)_w$ than $(S/N)_r$, presumably because the noise bias is more directly related to the signal-to-noise of the actual galaxy than to that of a counterfactual round version of the galaxy. Thus, the "noise bias" calibration in fact also calibrates the selection bias resulting from using $(S/N)_w$. This is therefore the appropriate S/N measure to use for selecting galaxies for the final IM3SHAPE catalog.

7.3. Shear Measurements with IM3SHAPE

IM3SHAPE is a maximum-likelihood model-fitting code, which in the mode used here fits de Vaucouleurs bulge and exponential disk components to galaxy images. It was described in Zuntz et al. (2013), where its performance on GREAT08 and its known biases were characterized.

We have slightly modified the model described therein, improving both its stability and its performance on the tests detailed in this paper (cf. §8). Previously each galaxy was modeled as the sum of two components, a bulge and a disk. In this paper each galaxy is modeled twice: once as a pure bulge and once as a pure disk. Whichever model has the higher likelihood is used in the catalog, unless that model is flagged as a bad fit (cf. §7.3.3). If both models are flagged the galaxy is excluded from the catalog.

The parameters of the best-fit model are found using the numerical optimizer LEVMAR[13] (Lourakis 2004), which implements the Levenberg-Marquardt algorithm (Levenberg 1944; Marquardt 1963), iterating towards a model image which minimizes the χ^2 with the data image.

For the SV data IM3SHAPE was run on r-band images only. For future data we plan to test fitting multiple bands simultaneously with marginalized relative amplitudes.

Since Zuntz et al. (2013), a number of additions and modifications have been made to the code; in this section we briefly review the code and its methodology, with particular focus on these changes.

The complete code with all the changes described below is available for download[14]. One particularly useful infrastructure improvement has been the implementation of a Python interface to the existing C functions. We use the Python interface to load data from MEDS files (via the `meds` module), select exposures, mask images, and compute most of the diagnostic information described in §7.3.3 below.

[13] `http://users.ics.forth.gr/~lourakis/levmar/`

[14] `https://bitbucket.org/joezuntz/im3shape`

The biggest change made to IM3SHAPE is the addition of a new model which fits multiple exposures of the same galaxy simultaneously. We now define our model parameters in a celestial coordinate system: a local tangent plane centered at the nominal Right Ascension and Declination of the galaxy. This model is then constrained by the pixel data from each epoch where the galaxy was observed, as discussed in §7.1.

7.3.1. Bulge or disk model

Each galaxy model is defined by six varied parameters: the amplitudes of either the bulge or disk components (A_b, A_d), a centroid relative to the nominal detection position ($\delta u, \delta v$), an ellipticity (e_1, e_2), and a half-light radius (r).

To compute the likelihood of a particular model for a given galaxy observed on a number of individual exposures, we use the local affine approximate of the WCS for each postage stamp (stored in the MEDS file) to transform these parameters into each image's local pixel coordinate system. Schematically,

$$\begin{aligned}&\{\delta u, \delta v, e_1^c, e_2^c, r^c\} \rightarrow \\ &\quad \big[\{\delta x, \delta y, e_1^p, e_2^p, r^p\}_{\text{Image 1}}, \\ &\quad \{\delta x, \delta y, e_1^p, e_2^p, r^p\}_{\text{Image 2}}, ...\big]\end{aligned} \tag{7-5}$$

where ^{c} indicates the parameters in celestial coordinates (u, v) and ^{p} indicates the transformed parameters in pixel coordinates (x, y). The amplitudes do not require any transformation, since the MEDS files have already put the postage stamps on the same photometric system.

Once we have the appropriate parameters for each postage stamp, we then build the galaxy models in pixel coordinates, each convolved by the correct PSF for that stamp, and compute a χ^2 of the model relative to the data, using the correct pixel noise. The total χ^2 from all the postage stamps then gives us the final likelihood to use for that set of model parameters. We can then iterate to find the maximum likelihood parameters for each galaxy. The maximum likelihood is typically found in less than 50 iterations. At 150 iterations, we stop the algorithm and declare failure.

The Levenberg-Marquadt code that we use to find the maximum likelihood, LEVMAR, does not directly handle problems where different weights are applied to each data point. The straightforward fix for this is to scale both the observed intensity I_p and the model m_p by the standard deviation of the intensity σ_p before passing them to LEVMAR:

$$\begin{aligned} I_p^{\text{LM}} &= I_p/\sigma_p \\ m_p^{\text{LM}} &= m_p/\sigma_p. \end{aligned} \tag{7-6}$$

This maintains the χ^2 per pixel that the LEVMAR algorithm uses as its objective function. The estimates of σ_p come from the weight map (as σ_p^{-2}) provided with the images.

7.3.2. Shear calibration

A significant problem with maximum likelihood shear estimators is that the peak of the likelihood distribution is not an unbiased estimator of the shear in the presence of noise (Refregier et al. 2012; Kacprzak et al. 2012). The fitted model parameters are a nonlinear function of pixel intensities affected by Gaussian noise, resulting in *noise bias* in the estimated shear values. The IM3SHAPE algorithm, being a maximum likelihood estimator, is known to suffer from this effect.

In addition, we find a small *selection bias*, which is introduced by using recommended IM3SHAPE flags (cf. §7.3.3) and the selection based on galaxy size and S/N (cf. §9.1). We also expect a small amount of *model bias* due to realistic galaxies not always being well fit by our bulge-or-disk model. This model bias is expected to be small compared to the requirements (Kacprzak et al. 2014).

To account for all of these sources of error in our shape measurements, we calculate bias corrections of the form shown in equation 3-4. Specifically, we fit for m and α as functions of $(S/N)_w$ (defined in equation 7-3) and R_{gp}/R_p (the FWHM of the PSF-convolved galaxy divided by the FWHM of the PSF) on simulated data from the GREAT-DES simulation (cf. §6.1). We run IM3SHAPE on the simulated data in the same way as we do on the DES data, including the same choices of input parameters.

In principle, the two multiplicative terms, m_1 and m_2 should be treated as independent biases. In practice, however, when averaged over many galaxies we find virtually no difference between the two. As such, we correct both e_1 and e_2 by the average $m = (m_1 + m_2)/2$.

We fit both m and α as two-dimensional surfaces in the S/N and size parameters. Due to the complicated structure of this surface, we fit m with 15 terms of the form $(S/N)_w^{-x}(R_{gp}/R_p)^{-y}$, where x and y are various powers ranging from 1.5 to 4. To control overfitting, we use a regularization term in the least-square fit and

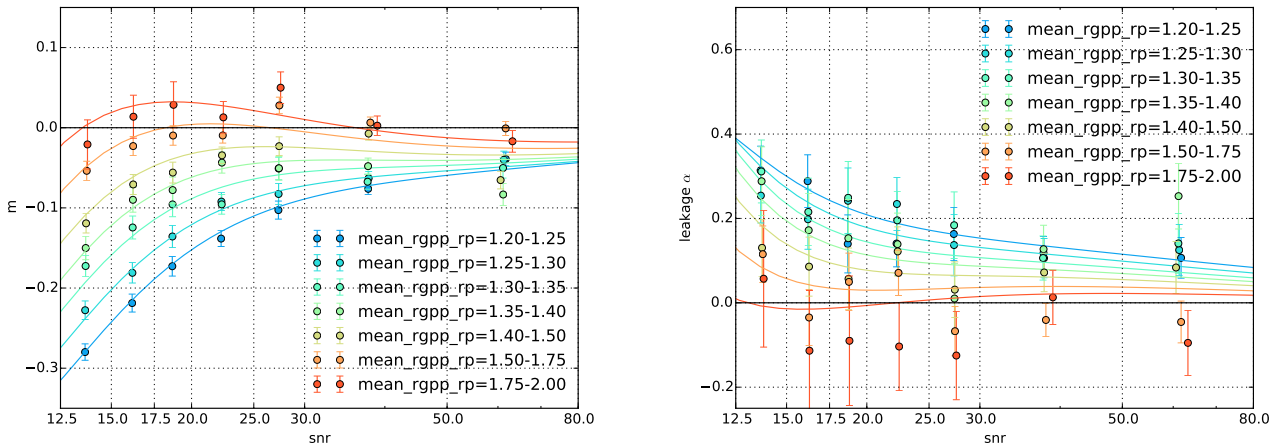

Fig. 12.— Shear bias for IM3SHAPE measurements on the GREAT-DES simulation: multiplicative bias (left) and PSF leakage (right), as functions of the measured $(S/N)_w$ and R_{gp}/R_p. The fits, which are used to calibrate the shear estimates on the data, are smooth functions in both of these variables. Solid lines show the fits vs $(S/N)_w$ at particular choices of R_{gp}/R_p.

optimize it such that the fitted surface has a reduced $\chi^2 = 1$. A similar procedure is applied to α, where we use 18 parameters in the fit. Figure 12 shows these fits as curves in $(S/N)_w$ in bins of R_{gp}/R_p. However, the actual functions are smooth in both parameters.

We check if our calibration is robust to the details of this model by (1) varying the number of terms in the basis expansion and (2) splitting the training data into halves. For both tests the changes in the mean multiplicative and additive corrections applied to the SV data do not vary by more than 1%.

In §7.2, we mentioned that $(S/N)_w$ is a biased measure of S/N with respect to shear, so if it is used to select a population of galaxies, it will induce a selection bias on the mean shear. R_{gp}/R_p similarly induces such a bias. Thus, when we bin the shears by these quantities to construct the calibration functions, there is a selection bias induced in every bin. The scale of selection bias reaches $m \simeq -0.05$ for the most populous bins. This is not a problem for the correction scheme so long as the overall selection is also made using these same quantities. In that case, the shear calibration automatically accounts for the selection bias in addition to the noise bias.

We tried using $(S/N)_r$ in the calibration model rather than $(S/N)_w$ to help reduce the level of the selection bias in each bin, but we found that it does not perform as well as using the standard $(S/N)_w$. Perhaps not surprisingly, the noise bias seems to be more related to the S/N of the actual galaxy than it is to the counterfactual round version of the galaxy used for $(S/N)_r$. In future work, it would be interesting to seek an effective shear calibration scheme that disentangles noise and selection biases, but we have not found one yet.

We use these fits to estimate the multiplicative and additive corrections to use for every galaxy in the IM3SHAPE catalog. However, it should be stressed that this bias estimate is itself a noisy quantity, being based on noisy estimates of the size and S/N. Therefore one should not directly apply the correction to each galaxy individually. Rather, the mean shear of an ensemble of galaxies should be corrected by the mean shear bias correction of that same ensemble (cf. §9.2).

Note that selection bias can appear whenever a subset of galaxies is selected from a larger sample. In the cosmological analysis, we apply recommended IM3SHAPE flags, cut on R_{gp}/R_p and $(S/N)_w$, and then typically split the galaxies into redshift bins. The redshift selection in particular is not used in the shear calibration process, so it is possible for there to be uncorrected selection biases in the different redshift bins. In §8.5, we test that the shear calibration nonetheless performs well in this scenario by applying the same selection procedure to the GREAT-DES simulation. There we demonstrate that all biases are removed to the required tolerance level in all redshift bins.

7.3.3. *Diagnostics*

After performing the shape measurement, we generate a large suite of diagnostic information based on the results of the fits to help identify objects that potentially should not be used for weak lensing. Many objects show evidence of imaging artifacts or some other problem that violates the assumptions we have made in the model, so we want to be able to remove these objects from the final shear catalog.

We distinguish two types of flags: "error" flags, which identify objects that should definitely be removed from any analysis, and "info" flags, which identify objects that may be somewhat contaminated, but which may have some value depending on the science application. Most of the info flags are derived by examining histograms of the relevant parameters and flagging extreme tails.

The full listing of IM3SHAPE flags is given in Appendix B. In §9.1, we will detail the final selection criteria that we recommend for the IM3SHAPE catalog, which will include both `error_flag==0` and `info_flag==0`. Moving to a less restrictive selection should only be done after carefully testing for the possibility of increased systematic errors.

7.3.4. *Galaxy weights*

We assign a weight to each shear measured by IM3SHAPE based on an estimate of the total shear uncertainty including both shape noise $\sigma_{\rm SN}$ (the standard deviation of the intrinsic ellipticities) and measurement uncertainty σ_e:

$$w = \frac{1}{\sigma_{\rm SN}^2 + \sigma_e^2}. \tag{7-7}$$

To estimate the appropriate weight for each galaxy, we use the measured shears from the GREAT-DES simulation. We group galaxies in bins of $(S/N)_w$ and R_{gp}/R_p. We then measure the width of the distribution of ellipticities in each bin, both by fitting a Gaussian to a histogram of the distribution and by measuring the sample variance directly. The larger of the two variance estimates is taken, and the weight is then given by the inverse variance.

We also impose a maximum weight set by the mean variance of all high-S/N bins. Otherwise spuriously low variance estimates in some sparsely populated bins result in very high weight values for those bins.

7.4. Shear Measurements with NGMIX

The code NGMIX is a general tool for fitting models to astronomical images (Sheldon 2014). The code is free software[15], and is available for download[16].

Both the PSF profile and the galaxy are modeled using mixtures of Gaussians, from which the name NGMIX is derived. Convolutions are thus performed analytically, resulting in fast model generation as compared to methods that perform the convolution in Fourier space.

7.4.1. *Exponential disk model*

For the galaxy model, NGMIX supports various options including exponential disks, de Vaucouleurs profiles (de Vaucouleurs 1948), and Sérsic profiles (Sérsic 1963), all of which are implemented approximately as a sum of Gaussians using the fits from Hogg & Lang (2013). Additionally, any number of Gaussians can be fit, either completely free or constrained to be co-centric and co-elliptical. For the DES SV catalogs, we used the exponential disk model.

Using this simple disk model results in detectable model bias (cf. §8.5). In simulations, we found this model bias was reduced when using a more flexible model, but the more flexible model was not implemented for real survey data in time for this release. We will explore improved modeling in detail for future DES analyses.

We construct the model in celestial coordinates and fit it to multiple epochs and bands simultaneously (cf. §7.1). The center, size and ellipticity are the same for all bands and epochs, but the flux is allowed to vary between bands. For this study we combined bands r, i, z, resulting in eight free parameters:

- u_c, v_c, the object center in celestial coordinates, relative to the fiducial center from the coadd object catalog. The units are arcseconds.
- e_1, e_2, the ellipticity.
- T, the area of the object, defined in terms of the unweighted moments of the Gaussian mixture $T = \langle x^2 \rangle + \langle y^2 \rangle$. The units are arcseconds squared.
- F_k, the flux in each of the r, i, z bands.

[15] `https://www.gnu.org/philosophy/free-sw.html`
[16] `https://github.com/esheldon/ngmix`

7.4.2. *Image fitting*

The NGMIX code supports multiple paradigms, all of which are used in the current analysis:

- Exploration of the full likelihood surface for a given set of model parameters with a Markov Chain Monte Carlo (MCMC) scheme, using either the standard Metropolis-Hastings algorithm (MH; Metropolis et al. 1953) or the recently introduced affine invariant method (Goodman & Weare 2010; Foreman-Mackey et al. 2013). The model can be fit directly to the pixel data, or it can include convolution by a point-spread function (PSF).

- Maximum-likelihood fitting using any of a variety of function minimizers. We used Levenberg-Marquardt (LM; Levenberg 1944; Marquardt 1963) as well as the method of Nelder & Mead (1965) (NM) in this work. The model can be fit directly to the pixel data, or it can include convolution by a PSF.

- Expectation Maximization (EM; Dempster et al. 1977), fitting directly to the pixels only. This method is used for PSF fitting.

For PSF measurement, the EM code is recommended, using an arbitrary number of completely free Gaussians. For DES we use three. EM is a good choice for PSF measurement, since it is extremely stable even with many components. By allowing all components to be completely free, the off-center PSF components that are occasionally found in the SV PSF images are properly fitted.

We choose to handle the WCS information by projecting each pixel into celestial coordinates and building both the galaxy and PSF models in that coordinate system.

Our procedure for fitting the galaxy shapes involves a number of steps:[17]

1. Estimate a flux for the object by fitting the PSF model to the galaxy with a single free parameter, which is the overall normalization (keeping the centroid fixed at its fiducial value).

2. Run NM to find the maximum likelihood model, guessing the flux from the result of step 1, and guessing the size to be the typical seeing size. We find NM to be more robust than LM for this fit.

3. Run LM starting from the maximum likelihood model to estimate the covariance matrix, since NM does not produce one. Relatively few evaluations are made in this step.

4. Run an MCMC chain with MH using the maximum likelihood position as a starting guess and the covariance matrix as a proposal distribution. We run a few thousand burn-in steps, followed by a few thousand post-burn-in evaluations. If the acceptance rate is outside the range $[0.4, 0.6]$ we reset the proposal distribution based on the covariance matrix from previous MH run, and run a new burn-in and post-burn in. If the acceptance rate remains outside the desired range we try again up to four times. These bounds on the acceptance rate are somewhat arbitrary, but for our problem we found that rates above 0.6 result in highly correlated chains, and lower than 0.4 can result in a poorly sampled peak.

7.4.3. *Shear estimation*

Multiple methods are supported for shear measurement, but for this study we adopted the "LENSFIT"-style method, based on the work of Miller et al. (2007). We found our implementation of this method to be sufficiently accurate for the precision of our current data set; for this study NGMIX measurements were instead limited by the use of an overly simple exponential disk model for galaxies (cf. §8.5).

The LENSFIT method involves multiplication by a prior on the distribution of galaxy ellipticities when estimating the expectation value of the ellipticity for each galaxy.

$$\langle e_\mu \rangle = \frac{\int \mathcal{L}(\boldsymbol{e}) p(e) e_\mu \mathrm{d}\boldsymbol{e}}{\int \mathcal{L}(\boldsymbol{e}) p(e) \mathrm{d}\boldsymbol{e}} \tag{7-8}$$

$$\simeq \frac{\sum_j p(e^j) e^j_\mu}{\sum_j p(e^j)}, \tag{7-9}$$

where $\mathcal{L}(\boldsymbol{e})$ is the likelihood and $p(e)$ is the prior on the galaxy shapes. We approximate the integral over the likelihood with the sum of points from an MCMC chain. The index μ takes values 1,2 for each ellipticity component such that $\boldsymbol{e} = (e_1, e_2)$; the ellipticity magnitude is given by e.

[17] We tried using the affine invariant fitter, and found it to be very robust, but the burn-in period was too slow for large-scale processing. This hybrid approach using both maximum-likelihood fitters and MH is significantly faster and sufficiently accurate.

Multiplying by an ellipticity prior reduces the effects of noise, which broadens and distorts the likelihood surface. However, application of the prior also biases the recovered shear, in effect reducing the "sensitivity" of the shear estimate. Miller et al. (2007) derive a measure of the sensitivity of this estimator to a shear $\boldsymbol{g}$, which is approximately given for each component by:

$$s_\mu \equiv \frac{\partial \langle e_\mu \rangle}{\partial g_\mu} \simeq 1 - \left[\frac{\int \left(\langle e_\mu \rangle - e_\mu \right) \mathcal{L}(\boldsymbol{e}) \frac{\partial p}{\partial e_\mu} \mathrm{d}e}{\int \mathcal{L}(\boldsymbol{e}) p(e) \mathrm{d}e} \right] \simeq 1 - \left[\frac{\sum_j \left(\langle e_\mu \rangle - e_\mu^j \right) \frac{\partial p}{\partial e_\mu}}{\sum_j p(e^j)} \right]. \tag{7-10}$$

No expression was formally derived by Miller et al. (2007) for the mean of the shear field acting on an ensemble of galaxies; however, it was proposed to use the same formula as derived for a constant applied shear:

$$g_\mu = \frac{\sum_i \langle e_\mu^i \rangle}{\sum_i s_\mu^i}, \tag{7-11}$$

where the index i runs over all galaxies in the measurement. In practice we also apply weights in both sums,

$$w = \frac{1}{2\sigma_{\rm SN}^2 + C_{1,1} + C_{2,2}}, \tag{7-12}$$

where $\sigma_{\rm SN}$ is the shape noise per component, which we have calculated to be 0.22 based on fits to COSMOS galaxies (cf. §7.4.4), and $C_{i,j}$ are elements of the 2x2 ellipticity subset of the covariance matrix produced by NGMIX.

The sensitivities in equation 7-10 do not transform as polarizations. Thus for practical shear measurements, such as tangential shear or two-point functions, which require rotation of the ellipticities, we choose to use a scalar sensitivity for each galaxy that is the mean of its two components.

7.4.4. Ellipticity prior

The LENSFIT method requires as input a prior on the shapes of galaxies, $p(e)$. The prior must be continuous for e_1, e_2 in the unit circle in order to evaluate the derivatives in equation 7-10.

In simulations, we found that the accuracy of the shear recovery was sensitive to the details of the ellipticity prior. For example, we ran the shear code on the GREAT-DES simulation presented in §6.1 using a prior with intrinsic variance in ellipticity 35% higher than the true variance, and found the multiplicative bias increased by $(1.3 \pm 0.2)\%$.

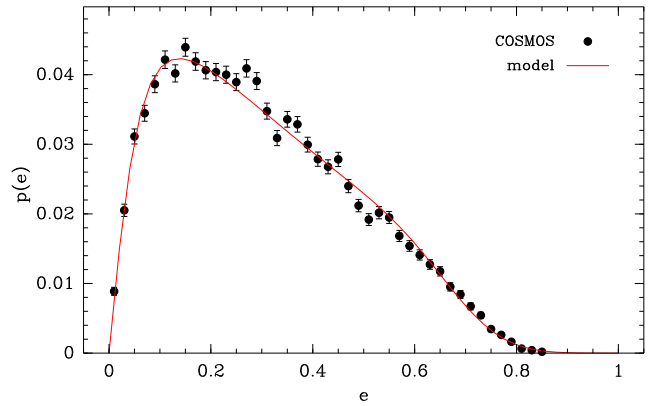

Fig. 13.— Distribution of shapes for COSMOS galaxies, selected as described in the text. The model fit was used as a prior for the NGMIX shear analysis.

For the data, we base our prior on the ellipticities of Sérsic model fits to COSMOS galaxies, as released by the GREAT3 team (Mandelbaum et al. 2014). We fit the observed distribution to a simple model

$$p(e) = 2\pi e A \, \frac{\left(1 - \exp\left[\frac{e-1}{a}\right]\right)}{(1+e)\sqrt{e^2 + e_0^2}} \, c(e) \tag{7-13}$$

$$c(e) = \frac{1}{2}\left(1 + \mathrm{erf}\left[\frac{e_{cut} - e}{\sigma_e}\right]\right). \tag{7-14}$$

This model is a modified version of that introduced in Miller et al. (2013). Note in particular the cutoff at high ellipticities achieved by using an error function. We find this formula improves the fit to the distribution of ellipticities.

This model is fit to the ellipticities of COSMOS galaxies selected to fall in a range of size and flux that corresponds to the galaxy population seen in our data. A comparison between the measured $p(e)$ and the fit model is shown in Figure 13. The best fit parameters are given in Table 1.

We use this same prior for all galaxies, but the distribution of COSMOS galaxy shapes depends on redshift. However, for this study the uncertainties due to the redshift dependence of the shape distribution were sub-dominant to model bias for NGMIX (cf. §8.5).

7.5. Blinding

Shape catalogs from both pipelines were blinded before they were used for any tests or SV science papers. This was to prevent the *experimenter bias* effect, wherein researchers work harder on finding bugs,

Parameter	Fit Value
A	0.025 ± 0.002
a	2.0 ± 0.2
e_0	0.079 ± 0.003
e_{cut}	0.706 ± 0.004
σ_e	0.125 ± 0.006

Table 1: Parameters for the ellipticity prior used with the NGMIX shear code, with best fit values from fitting to the distribution of shapes of COSMOS galaxies.

tuning methodology, etc. when results are inconsistent with a previous experiment, or otherwise do not match expectations, than when they do match (cf. Klein & Roodman 2005).

SV shear catalogs were blinded by scaling all measured shears by a secret factor generated by an algorithmic but unpredictable process (using an MD5 hash of a code phrase) to be between 0.9 and 1.0. This unknown scaling meant that it was harder for DES members to, for example, accidentally tune results to get the σ_8 value predicted by Planck. Only once the analysis for a given paper was finalized has the unblinded catalog been made available.

This was a gentle blinding approach that was appropriate for the relatively loose statistical constraints that will come from SV data. It has the useful feature that, being linear, correlation tests on it such as those listed in this paper remain valid. It has a significant downside in that it is asymmetric - unblinding could only increase the measured σ_8, so the potential for bias was still present. We will consider new blinding methodologies for future data.

8. Tests of the Shear Measurements

We have developed an extensive test suite to check that the shear catalogs do not have significant systematic errors that would adversely affect weak lensing science. While there is no way to definitely prove that the shear catalogs are free of all possible systematic errors, there are many tests that can reveal systematic errors that might be present in the data. These tests are formulated as "null tests", which should have zero signal in the absence of systematic errors. Most of our null tests are similar to ones that have been performed in previous analyses (cf. e.g. Jarvis et al. 2003; Schrabback et al. 2010; Velander et al. 2011; Heymans et al. 2012b; Jee et al. 2013; Kuijken et al. 2015).

These null tests can be broken up into several broad categories.

1. **Spatial tests** check for systematic errors that are connected to the physical structure of the camera. Examples of these would be errors in the WCS correction, including effects like edge distortions or tree rings (Plazas et al. 2014), and errors related to features on the CCDs such as the tape bumps. (§8.1)

2. **PSF tests** check for systematic errors that are connected to the PSF correction. This includes errors due to inaccurate PSF modeling as well as leakage of the PSF shapes into the galaxy shape estimates. (§8.2)

3. **Galaxy property tests** check for errors in the shear measurement algorithm related to properties of the galaxy or its image. This can include effects of masking as well, which involve the other objects near the galaxy being measured. (§8.3)

4. **B-mode statistics** check for systematic errors that show up as a B-mode signal in the shear pattern. The gravitational lensing signal is expected to be essentially pure E-mode. Most systematic errors, in contrast, affect the E- and B-mode approximately equally, so the B-mode is a sensitive test of systematic errors. (§8.4)

5. **Calibration tests** check for systematic errors that affect the overall calibration of the shears. If all of the shear values are scaled by a constant factor, most null tests remain zero (if they were zero to start with). Furthermore, there are no known absolute shear calibration sources that we can use to calibrate our results. For these reasons it can be hard to tease out errors in the calibration from the data. However, we can use simulated data where the true shear is known to check that we recover the correct values. (§8.5)

6. **Cross-catalog comparisons** check that the two shear catalogs are consistent with each other. Because we have two shear catalogs available for testing, we can check that the two give consistent results, thus potentially uncovering problems that may be in one shear catalog but not the other (or have different levels in each). Considering the large differences between the NGMIX and IM3SHAPE codes, these are very non-trivial tests. (§8.6)

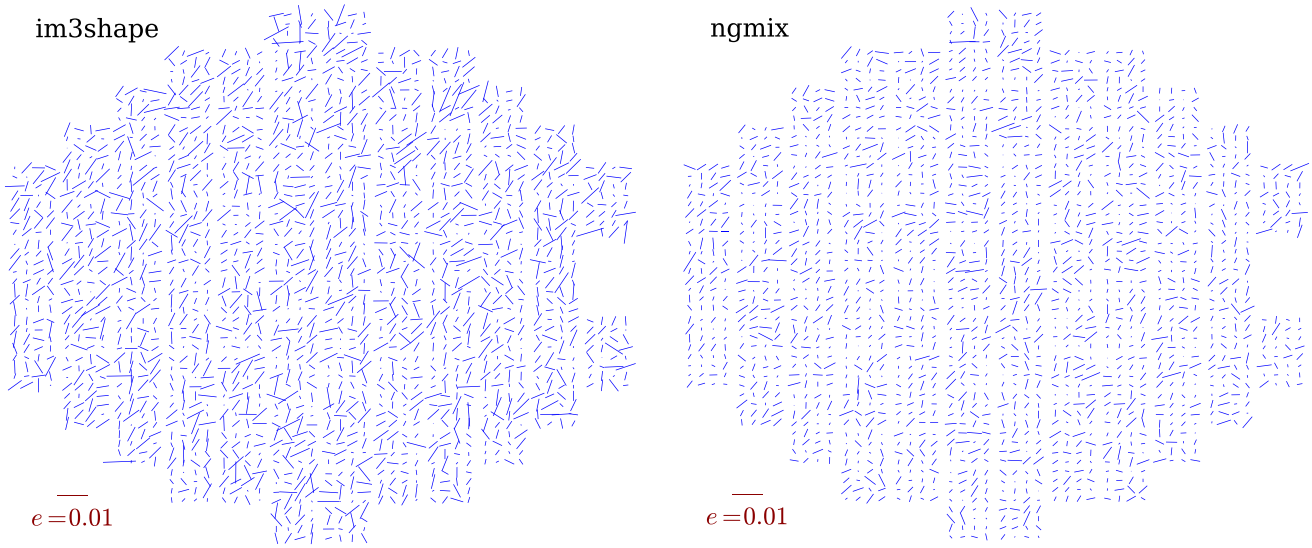

Fig. 14.— Whisker plots of the mean shear binned by position in the focal plane for IM3SHAPE (left) and NGMIX (right). As in Figure 8, the length of each whisker is proportional to the mean shear, and the orientation is aligned with the direction of the ellipticity. A 1% whisker is shown for scale in the lower left.

One caveat to keep in mind with the various null tests is that we do not necessarily expect the overall mean shear to be precisely zero. The SV region is small enough that the rms value of the mean shear due to cosmic variance is expected to be about 4×10^{-4}. In fact, the overall mean shear is measured to be

$$\text{IM3SHAPE} \quad \begin{array}{l} \langle e_1 \rangle = 0.1 \times 10^{-4} \\ \langle e_2 \rangle = 6.8 \times 10^{-4} \end{array}$$

$$\text{NGMIX} \quad \begin{array}{l} \langle e_1 \rangle = -0.4 \times 10^{-4} \\ \langle e_2 \rangle = 10.2 \times 10^{-4}. \end{array}$$

These are both about 2σ from zero given the expected cosmic variance, so it may be due to an additive systematic error affecting both codes. However, the fact that they roughly agree with each other suggests at least the possibility that it could be a real cosmic shear signal. In any case, each of the null tests look for variations *relative* to this overall mean shear to find dependencies that may indicate systematic errors.

In §8.1 – §8.6 we show the results of our null tests in each of the above categories. Then §8.7 summarizes these results and tries to quantify the total possible systematic errors that may be present in the shear catalogs.

8.1. Spatial Tests

There are many potential sources of systematic error related to the camera and telescope optics that can cause a spatial dependence of the shear with respect to the camera's field of view. The telescope distortion pattern and some of the optical aberrations are essentially static in time. The CCDs have bad columns and other defects, including the tape bumps mentioned in §4.2. There are also distortion effects at the CCD edges due to the electric field lines becoming non-parallel as well as tree ring distortion patterns due to doping variations in the silicon (Plazas et al. 2014).

8.1.1. Position in the field of view

To check that we have adequately corrected for effects that are connected with the telescope and camera, or that they are small enough to ignore, we bin the shear spatially with respect to the field of view.

Figure 14 shows the mean shear as a function of position on the focal plane for both IM3SHAPE (left) and NGMIX (right). Each whisker is the mean shear of all galaxies that were ever observed in that area of the focal plane. As our shear measurements uses information from multiple epochs, each measurement contributes to this plot multiple times: once for each single-epoch observation of that galaxy.

This figure is similar to Figure 8, which showed the residual PSF pattern as a function of position on the focal plane. These plots are a noisier due to the shape noise of the galaxies, but there is a hint of the same radial patterns that were seen for the PSF residuals, especially in the NGMIX results, which are slightly less noisy due to the higher number of galaxies in the cat-

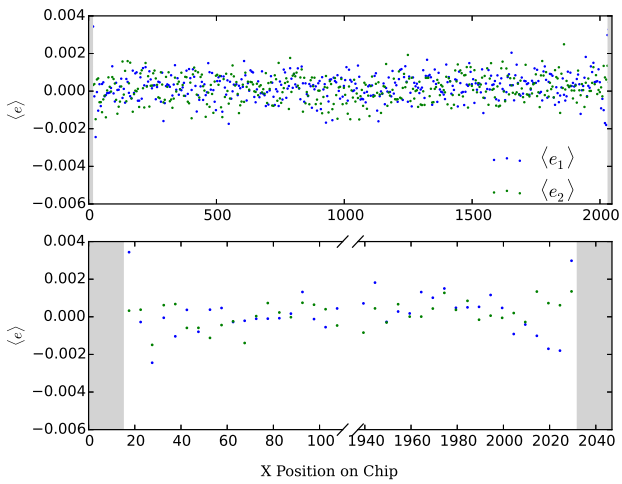

Fig. 15.— The mean shear $\langle e_1 \rangle$ (blue) and $\langle e_2 \rangle$ (green) for NGMIX, binned by column number X; X runs along the readout rows of the CCDs. (The corresponding plot for IM3SHAPE looks similar but noisier.) The bottom panel shows the same data blown up near the left and right edges to highlight the effect of the edge distortion. We mask the 15 columns along each edge where the distortion is strongest, but there is still a slight bias in the e_1 component of the shears up to 40 pixels from the edge.

alog. This is not surprising; we expected these PSF interpolation errors to leak into the galaxy shapes.

8.1.2. Position on CCD

If we bin the shears by their column number, irrespective of the CCD number, as shown in Figure 15, we can see the effect of something known as "edge distortion" (Plazas et al. 2014). This is where the electric field lines in the detector become slightly non-parallel near the edges of the CCDs. The cross section of the pixels becomes rectangular, elongated in the direction towards the edge of the CCD. Plazas et al. (2014, their Figure 6) showed that this effect led to photometric biases of ~ 20 mmag at ~ 30 pixels from the edge of the CCDs. Since flux and shape respond to the astrometric variation at the same order, this implies that we should expect shape residuals of about $\delta e_1 \sim 0.02$ near the edge of the CCDs.

Figure 15 shows the mean shape measured by NGMIX binned by column number. As we do not measure the single-epoch shape, any effect on the shapes has been reduced by a factor of about 10, the number of single-epoch exposures of each galaxy. So we might expect a signal of $\langle e \rangle \sim 0.002$. There does seem to be a slight effect visible in Figure 15 at this level for e_1, although it is not highly significant. The effect of the edges is even less evident when binning by the row number (not shown).

To quantify how much this edge effect might impact the overall shear signal, we estimate that the effect is only significant for about 20 pixels on any edge. This is a fraction of $40/2048 + 40/4096 = 0.015$ of the area. Galaxies have ~ 10 chances to fall in this area, so about 15% of the galaxies may have a spurious shear of ~ 0.002. The net additive systematic shear from this effect is thus about $c_{\rm rms} = 8 \times 10^{-4}$. This is well below the requirements on additive systematic errors given by equation 3-11, $c_{\rm rms} < 2 \times 10^{-3}$; however, it will not be below the expected requirements for DES 5-year data. Therefore, we plan to remove this effect directly in the astrometry solution in future DES data analyses.

8.1.3. Tangential shear around field centers

The telescope distortion pattern is approximately a fifth order radial function centered near the center of the field of view. If it is not corrected it can induce spurious shears oriented either radially or tangentially relative to the field centers. We look for this effect by measuring the tangential shear pattern around the set of field centers; essentially this is similar to a galaxy-galaxy lensing measurement where the telescope pointings play the role of the lenses.

Figure 16 shows the results of this test for both the IM3SHAPE and NGMIX shear measurement pipelines. Uncertainties are jackknife estimates, made by splitting the total area into 152 equal-area sub-fields. At large scales, the measurements are consistent with zero, but at scales less than about 10 arcminutes there are a few consecutive bins with $\sim 1\sigma$ deviation from zero in both cases. The IM3SHAPE results show a slight oscillating pattern, and the NGMIX results are slightly negative (a radial shear pattern).

None of these features are highly significant, especially since the points are somewhat correlated, so it may just be a noise fluctuation. Also, since the telescope distortion is largest at the edge of the field of view, we expected the absolute mis-estimation of the distortion to be largest at a separation of around 1 degree. Furthermore, IM3SHAPE and NGMIX use exactly the same WCS solution, since it is incorporated into the MEDS files directly. So the fact the tangential shear patterns are different in the two cases, and most significant near the center, indicates that this is probably not

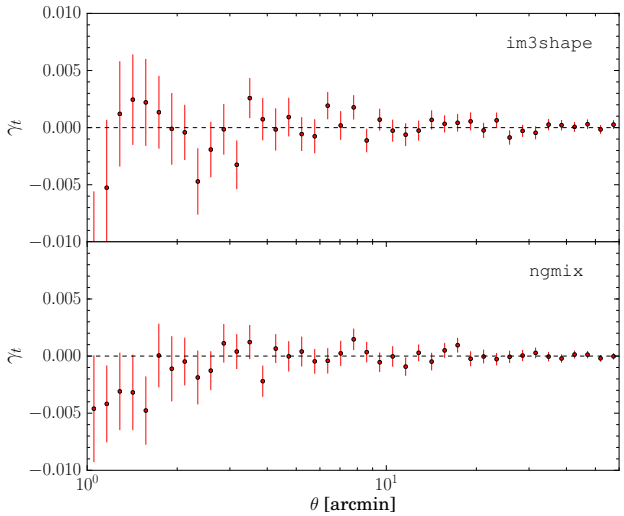

Fig. 16.— The tangential shear of galaxies in IM3SHAPE (top) and NGMIX (bottom) around field centers. Both measurements are approximately consistent with zero, although at scales less than 10 arcminutes, both show a slight departure from the expected null signal. The magnitude of this effect is well below our requirements in both cases.

due to errors in the WCS solution, although we do not have a good hypothesis for a plausible cause.

We can estimate the magnitude of this potential additive systematic error in the same manner as we used above for the edge distortions. The mean spurious shear in this case has a magnitude of at most 0.005 in both cases and occurs over a relative area of about $(2'/62')^2 = 0.001$. The net additive systematic shear from this effect is thus at most $c_{\rm rms} = 2 \times 10^{-4}$, well below our requirements for an additive systematic shear.

We also looked at the shear around the CCD corners. While there was a very slight hint of a non- zero signal at small scales, the magnitude is even smaller than the shear around the field centers.

8.2. PSF Tests

If the PSF interpolation is not sufficiently accurate or if the shear algorithm does not fully account for the effects of the PSF convolution, the resulting shear estimates will include a spurious additive error that is correlated with properties of the PSF.

We look for such additive errors by examining: (1) the mean shear binned by PSF ellipticity and PSF size, (2) the PSF-shear two-point correlation function and derived quantities, and (3) the tangential shear measured around stars.

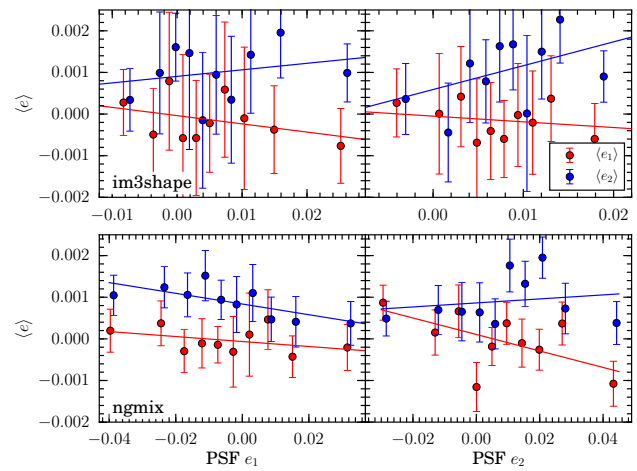

Fig. 17.— The mean galaxy shear as a function of the input PSF ellipticity (e_1 left, and e_2 right) for IM3SHAPE (top) and NGMIX (bottom). The solid lines show the best linear fit without binning. Note the range of the abscissa is different for the NGMIX and IM3SHAPE plots. The NGMIX measurements are averaged over r, i, and z-band images, while the IM3SHAPE measurements use r-band images, and different models are used for measuring the ellipticity, resulting in different PSF ellipticity ranges for the two catalogs.

8.2.1. *PSF leakage*

As we introduced in equation 3-4, we assume that a component of the additive bias in the shear estimates comes from imperfect correction of the PSF, resulting in a term proportional to the PSF shape:

$$e_{\rm gal} = e_{\rm true} + \alpha e_{\rm PSF} + c. \qquad (8\text{-}1)$$

A measured slope of galaxy ellipticity vs. PSF ellipticity can be identified as α, where we use the mean PSF shape over all epochs.

The mean shear as a function of PSF ellipticity is shown in Figure 17 for both IM3SHAPE and NGMIX galaxies. The points represent the mean galaxy ellipticity in each of 10 equal-number bins of PSF ellipticity. The line represents the best fit to the individual (unbinned) galaxy shapes. The slopes of the linear fits range from -2.0% to 5.7% $\pm$ 3% for IM3SHAPE and from -2.0% to 0.5% $\pm$ 1% for NGMIX. The slopes are consistent with no PSF leakage for both catalogs.

To obtain a more precise estimate of α, we compute the (weighted) average of the slopes of the red

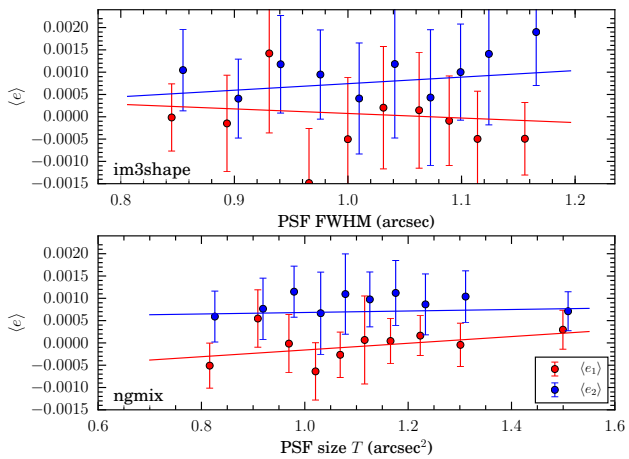

Fig. 18.— The mean galaxy shear as a function of the input PSF size for IM3SHAPE (top) and NGMIX (bottom). The solid lines show the best linear fit without binning.

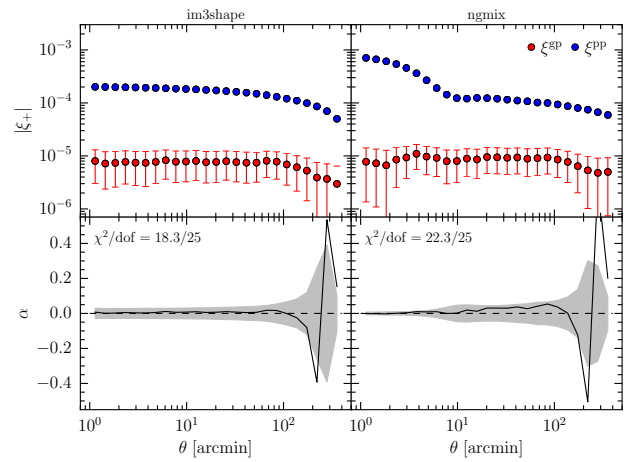

Fig. 19.— The calculation of the PSF leakage parameter α, which is given in equation 8-2. The top plots show ξ^{gp} (red), the cross correlation of the galaxy shapes with the PSF shapes, and ξ^{pp} (blue), the auto-correlation of the PSF shapes, for IM3SHAPE (left) and NGMIX (right). The bottom plots show α, which is a measure of the leakage of the PSF shapes into the galaxy shapes as a function of scale. The grey band shows the sample variance plus shape noise uncertainty for α. The $\chi^2/\mathrm{d.o.f.}$ is given for α over all scales.

lines on the left plots (i.e. $\langle e_1 \rangle$ vs. PSF e_1) and the blue lines on the right plots (i.e. $\langle e_2 \rangle$ vs. PSF e_2). For IM3SHAPE we find $\alpha = 0.008 \pm 0.025$, and for NGMIX $\alpha = -0.001 \pm 0.007$. There is no evidence for non-zero α; however, for IM3SHAPE, we cannot definitively confirm that $|\alpha| < 0.03$ (cf. equation 3-15) given the uncertainty in the estimate.

We similarly plot the mean shear as a function of PSF size in Figure 18 for both IM3SHAPE (left) and NGMIX (right). Linear best-fit lines are also included. The slopes here are also consistent with zero, being on the order of 0.1% or less, which indicates negligible dependence of the mean shear on the PSF size.

8.2.2. *Star/galaxy cross-correlation*

Another estimate of the leakage factor α comes from the cross-correlation of the galaxy shapes with the PSF shapes, ξ_+^{gp}. Writing ξ_+^{gp} in terms of equation 8-1 and solving for α, we find that

$$\alpha = \frac{\xi_+^{gp} - \langle e_{\rm gal} \rangle^* \langle e_{\rm PSF} \rangle}{\xi_+^{pp} - |\langle e_{\rm PSF} \rangle|^2}, \qquad (8\text{-}2)$$

where ξ_+^{pp} is the auto-correlation function of the PSF shapes, $e_{\rm PSF}$.

While this nominally gives us an estimate of α as a function of scale, α is not a scale-dependent quantity. It quantifies a point process, the possible leakage of the PSF shape into the galaxy shape estimates. Therefore, we expect this estimate of α to be consistent at all scales, given the uncertainties in the estimate.

The measured ξ_+^{gp} and ξ_+^{pp} correlation functions are shown in the top panels of Figure 19 for IM3SHAPE (left) and NGMIX (right). α is then calculated based on these and shown in the lower panels. Due to sample variance, α can be non-zero in this test even if the measured shears have no PSF contamination. We use the mock catalogs described in Becker et al. (2015) to compute the total uncertainty for α. These catalogs are populated with PSF shapes by using the PSF shape from the nearest observed galaxy to each mock galaxy. We then use the full suite of 126 mock catalogs to compute the total uncertainty on α including both shape noise and sample variance.

We find that both IM3SHAPE and NGMIX show no significant PSF contamination in this test, with a total $\chi^2/\mathrm{d.o.f.}$ of 18.3/25 and 22.3/25 for α computed over all scales. The best fit value for α in each case, properly taking into account the correlations (Avery 1996), is $\alpha = 0.010 \pm 0.023$ for IM3SHAPE and $\alpha = -0.008 \pm 0.006$ for NGMIX, both below the requirement of $|\alpha| < 0.03$ from equation 3-15, although in the case of IM3SHAPE we are only able to constrain $|\alpha|$ to be less than 0.03 at about 1σ.

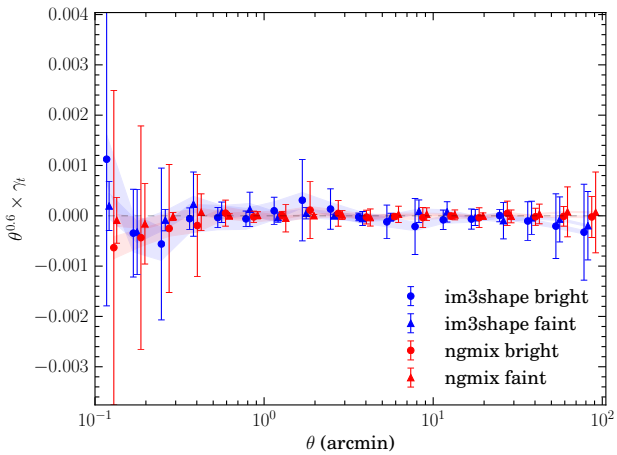

Fig. 20.— Tangential shear around stars for IM3SHAPE (blue) and NGMIX (red). Stars are split into to two bins of i-band magnitude, where "bright" means $14 < m_i < 18.3$ and "faint" means $18.3 < m_i < 22$. The faint sample includes stars used for PSF modeling; bright stars are excluded to avoid the brighter-fatter effect (cf. §4.2). Shaded regions represent 1σ shape noise uncertainty, while error bars are from jackknifing the stars.

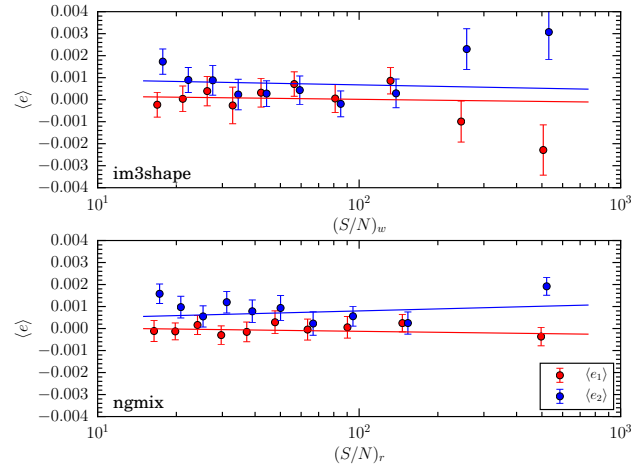

Fig. 21.— The mean galaxy shear as a function of the signal-to-noise for IM3SHAPE (top) and NGMIX (bottom). For IM3SHAPE we test against $(S/N)_w$ (cf. equation 7-3), which is one of the parameters used for the calibration, so it includes corrections for selection bias. For NGMIX, we test against $(S/N)_r$ (cf. equation 7-4), which does not induce any significant selection bias from the binning.

8.2.3. *Tangential shear around stars*

If the PSF correction is incomplete, there may also be a residual signal seen in the mean tangential shear around stars, which could potentially contaminate galaxy-galaxy lensing studies. To test for this, we measure the tangential shear around the positions of stars in both IM3SHAPE and NGMIX for "bright" ($14 < m_i < 18.3$) and "faint" ($18.3 < m_i < 22$) stellar populations. In all cases the signal, shown in Figure 20, is consistent with zero. The shape noise uncertainty is shown as the shaded regions. The error bars are jackknife uncertainty estimates.

The test using the faint stars primarily checks for effects related to PSF interpolation and PSF modeling. The bright stars are not themselves used to constrain the PSF model (cf. Figure 6), so these stars instead check for problems related to deblending and sky estimation errors in the outskirts of bright stellar halos. We see no evidence of any systematic errors around either set of stars.

8.3. Galaxy Property Tests

There are many properties of the galaxy images that should be independent of the shear, but which in practice can be correlated with systematic errors in the shear measurement. For example, some of the properties we tested during the course of our analysis were: the size of the postage stamp, the number of neighbors being masked, the fraction of the stamp area being masked, the estimated bulge-to-disk ratio, the galaxy's signal-to-noise, and the galaxy size. These were all extremely helpful diagnostic tools during the analysis, but here we only present the final two, which initially showed evidence of systematic errors and took the most effort to resolve.

Using S/N or the galaxy size for selections is quite natural, since estimating the shear for small, faint galaxies is more challenging than for large, bright galaxies. However, measurements of these quantities can be correlated with the galaxy ellipticity, and thus an applied shear. Binning the data for the null test by these properties can thus induce selection effects and produce a net mean spurious ellipticity. This was already discussed in §7.2 with respect to S/N. We need to do something similar to construct a proper null test for the galaxy size.

8.3.1. *Galaxy signal-to-noise*

The null test for checking that the galaxy shapes are independent of S/N requires different measures of S/N for each catalog. As described in §7.3.2, for IM3SHAPE we calibrate the bias in the shear measure-

ments from simulations as a function of $(S/N)_w$ and R_{gp}/R_p. As such, the selection bias that is induced by binning on $(S/N)_w$ (cf. §7.2) is accounted for as part of the calibration. Thus, the appropriate null test on the data is to check that the mean shear is independent of $(S/N)_w$, as shown in the top panel of Figure 21. There is no apparent bias in the mean shear down to $(S/N)_w = 15$.

NGMIX does not perform any calibration, so the null test for it requires a S/N measure that does not induce selection biases from the binning. For NGMIX, we use $(S/N)_r$ (cf. equation 7-4), which did not induce any selection biases when we tested it on simulated data. The bottom panel of Figure 21 shows that the mean estimated shear for NGMIX is independent of this "roundified" signal-to-noise measure down to $(S/N)_r = 15$.

Previous versions of the NGMIX catalog had shown a very significant bias in the lowest S/N bin in this plot before we realized that the bias was being induced by our galaxy selection from the cut on $(S/N)_w$. IM3SHAPE calibrates this kind of selection bias, but when that calibration is faulty, it too could show a bias in the lowest S/N bin. Figure 21 shows that the final catalogs do not have any such bias. The points are consistent with the mean value (which, as we mentioned, is not necessarily expected to be zero) and show maximal deviations less than our required $c_{\rm rms} < 2 \times 10^{-3}$ (equation 3-11).

8.3.2. *Galaxy size*

Similar considerations apply to the null tests for galaxy size. Since IM3SHAPE corrects for selection bias using the measured R_{gp}/R_p, this is the appropriate quantity to use for the null test regarding galaxy size. The top panel of Figure 22 shows the mean estimated shear binned by R_{gp}/R_p. The IM3SHAPE measurements show no evidence of any dependence of the shear estimates on the size of the galaxy.

For NGMIX, we need to use a size measure that is independent of the shape of the galaxy. The internal parameter that NGMIX uses for the size of the galaxy in its model is $T = I_{xx} + I_{yy}$, the standard second moment measure of the size of a galaxy; however, this quantity changes with applied shear. If a round galaxy is sheared by an area-preserving[18] shear g, then the measured size will be

$$T^{(g)} = T^{(g=0)} \left(\frac{1+|g|^2}{1-|g|^2} \right). \tag{8-3}$$

For non-round galaxies, an applied shear tends to make the estimated size T larger when the shear is aligned with the galaxy shape and smaller when it is anti-aligned. This can lead to an apparent bias in the measured shapes with respect to the measured value of T. If the mean PSF shape were precisely round, this bias should average out over an ensemble of galaxies; however, our PSFs have a preferred direction, which breaks the symmetry and leads to an apparent bias in the mean shape with respect to T.

In parallel to our definition of $(S/N)_r$ as the signal-to-noise that the galaxy *would have had if it were round*, we similarly construct an estimate of the size that the galaxy would have had if it were round:

$$T_r \equiv T \left(\frac{1-|e|^2}{1+|e|^2} \right), \tag{8-4}$$

where e is the estimated shape of the galaxy. Binning the shears by this quantity should thus not induce any selection bias from the binning itself. The lower panel of Figure 22 shows the results of this test for NGMIX. There is no apparent dependence of the mean shape on this "roundified" measure of the size of the galaxy.

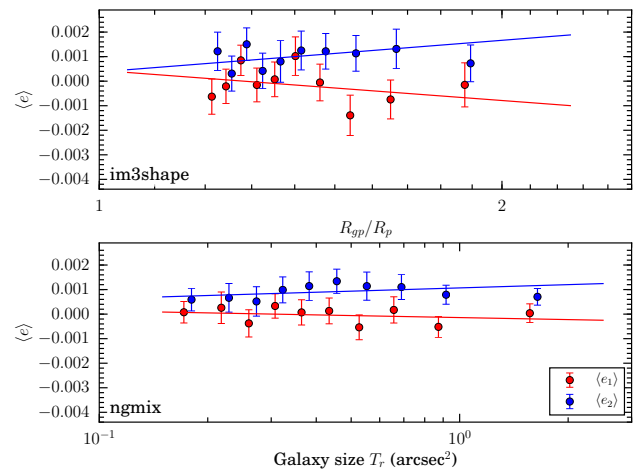

Fig. 22.— The mean galaxy shear as a function of the galaxy size for IM3SHAPE (top) and NGMIX (bottom). For IM3SHAPE we test against R_{gp}/R_p, which is one of the parameters used for the calibration, so it includes corrections for selection bias. For NGMIX, we test against a size measure, T_r (cf. equation 8-4), that does not induce and significant selection bias from the binning.

[18] By area-preserving, we mean that the determinant of the distortion matrix is unity: $A = \frac{1}{\sqrt{1-|g|^2}} \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}$.

In both cases the slopes are consistent with zero and show maximal deviations well below our required $c_{\rm rms} < 2 \times 10^{-3}$ (equation 3-11).

8.4. B-mode Statistics

The deflection field induced by gravitational lensing has the special property that it is essentially curl-free. Since this is also true of electric fields, the shear field is generally referred to as being an "E-mode" field. The corresponding divergence-free "B-mode" field can be considered as corresponding to an imaginary convergence (Schneider et al. 2002).

In fact gravity can induce a slight B-mode from source clustering (Schneider et al. 2002), multiple deflections (Krause & Hirata 2010), and intrinsic alignments (Crittenden et al. 2001). But in practice all of these effects are well below the level at which we could measure them, which means that any significant measured B-mode is almost certainly a sign of uncorrected systematic errors in the shears.

We calculate B-mode statistics of the shear field by computing linear combinations of the binned shear two-point correlation function values that are insensitive to any E-mode signals, modulo a very small amount of computable E- to B-mode leakage. See Becker (2013) for details. In this application, we have chosen linear combinations that approximate band-powers in Fourier space as described in Becker & Rozo (2014). Finally, we have used the mock catalogs described in Becker et al. (2015) to compute the shape noise and sample variance uncertainty for the statistics. These mock catalogs include the survey mask and match the shape noise and source photometric redshift distribution for each of the two shear catalogs. We have used 126 mock catalogs in total.

Figure 23 shows the measured B-mode for each catalog using the most conservative selection described below. Each band-power measurement is plotted at its central location in ℓ. Adjacent points are highly correlated and the χ^2 given in the figure accounts for the correlation. We find a $\chi^2/\mathrm{d.o.f.}$ of 22.3/20 for NGMIX and 16.1/20 for IM3SHAPE indicating no significant B-mode contamination in the shear field.

8.5. Calibration Tests

It is difficult to test the overall shear calibration using the data alone. However, we can use the GREAT-DES simulation described in §6.1 to test the performance of the two shear algorithms on relatively realistic images with known applied shear.

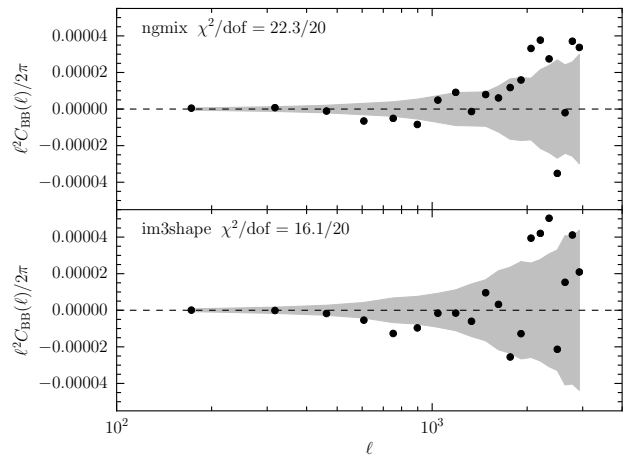

Fig. 23.— The measured B-mode for NGMIX (top) and IM3SHAPE (bottom). Each band power measurement is plotted at its central location in ℓ. The grey band shows the uncertainty on the measurement due to both sample variance and shape noise. Adjacent points are highly correlated and the indicated χ^2 accounts for the correlations.

Since IM3SHAPE uses this simulation to calibrate the shear measurements (cf. §7.3.2), the overall corrected shears should be accurate, almost by construction. The calibration is done without weighting, but here we use the same weights that we recommend for the data (cf. §7.3.4). The mean shear is thus not mathematically guaranteed to be exactly zero, but indeed the net bias after applying the calibrations is negligible: $m_1 = 0.0008 \pm 0.0015$ and $m_2 = -0.0068 \pm 0.0015$.

For NGMIX, the overall calibration error is a more relevant test. The priors used for GREAT-DES were the same as used for the DES SV data, which is expected to be appropriate given the general agreement between the galaxy properties in the simulation and the data (cf. §6.1). The overall calibration error for NGMIX is measured to be $m_1 = -0.030 \pm 0.0015$ and $m_2 = -0.035 \pm 0.0015$. This is not quite meeting our requirement of $|m| < 0.03$ from equation 3-10.

Considering that many science applications will use tomography to investigate the evolution of the shear signal with redshift, it is interesting to look at the calibration of both shear codes as a function of redshift. We use the known photometric redshifts of the galaxies from the COSMOS data where the galaxy images used for the simulation originated to test whether the calibration is robust to different distributions of galaxy properties in general, and as they vary with redshift in

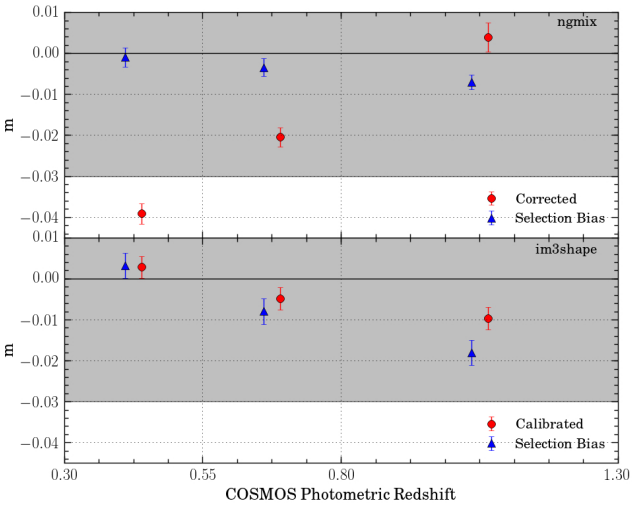

Fig. 24.— Multiplicative shear bias in three bins of photometric redshift for NGMIX (top) and IM3SHAPE (bottom), as calculated using the GREAT-DES simulation. In both cases, the same selection and weights were used as for the real data. The red circles denote the average bias in each bin after correcting for the sensitivity (NGMIX) or the calibration (IM3SHAPE). The blue triangles show an estimate of selection bias, calculated using the known true ellipticities. The grey band in both panels marks the $\pm 3\%$ requirement for SV data.

particular. This also tests that the tomographic selection process itself does not lead to significant selection biases for either catalog.

Figure 24 shows the results of performing this test for NGMIX (top) and IM3SHAPE (bottom), taking galaxies in different ranges of photometric redshift, based on the redshift bins that will be used for the cosmology constraints (The Dark Energy Survey Collaboration et al. 2015). Note that the redshift information was not used in the calibration process for IM3SHAPE, so the variation with redshift is a non-trivial test of the correction. Prior to calibration, we find a significant bias in each of the three redshift bins, $m = (-0.039, -0.058, -0.072)$. After calibration (red circles, Figure 24), the net multiplicative bias for IM3SHAPE is reduced to a level well within the requirements. This indicates that the derived corrections are robust to galaxy selections based on redshift.

We also tested the performance of IM3SHAPE's PSF leakage calibration as a function of redshift (not shown). The overall leakage before calibration is $\alpha = (0.070, 0.112, 0.102)$ for the three redshift bins. After calibration, $\alpha = (0.001, 0.021, -0.005)$, which demonstrates good performance of the leakage calibration as well.

The bias for NGMIX is seen to be outside of the requirement band for the lowest redshift bin, and then rise to acceptable levels in the two higher bins. We believe this is because the proportion of bulge galaxies is highest at low redshift, and the NGMIX exponential model has significant model bias for these galaxies. As the proportion of bulges decreases at higher redshift, the mean model bias decreases, and the calibration is within our requirements.

To test the hypothesis that we are measuring a model bias for the exponential disk model, we implemented a more flexible model and applied it to this simulation. This model is a simple two-component bulge and disk model, where the bulge fraction is determined not by a simultaneous fit with other parameters but by an initial comparison of two separate bulge and disk fits to the galaxy image. A similar model used in the Sloan Digital Sky Survey for galaxy fluxes was known as the "composite" model (Abazajian et al. 2004). For the composite model we find biases $\sim 1\%$, suggesting that the larger bias evident for the exponential model is principally model bias. Unfortunately, we were unable to apply this new composite model to the DES data in time to be used for this paper.

The blue triangles in Figure 24 show the estimated selection bias in each bin induced by our various selection criteria. We take the COSMOS shape estimates (Kannawadi et al. 2015) of the galaxies used for the simulation, shear them by our known applied shear, and then apply the same selection criteria we use for each of the two algorithms. The selection bias from the IM3SHAPE cuts is at most 2% in the highest redshift bin, which is largely corrected by the calibration scheme. The selection bias for NGMIX is less than 1% for all redshift bins.

8.6. Cross-catalog Comparisons

Another powerful test is to compare the two independent shear catalogs, IM3SHAPE and NGMIX. We use two very different strategies when generating these catalogs. For IM3SHAPE we use simulations to determine the shear calibration, and corrections are applied to the shear measurements on real data. For NGMIX we expect relatively little noise bias, but the sensitivity of the shear estimator has to be calculated from the data and applied to the shear measurements. We do expect some model bias for NGMIX. Furthermore, the PSF is treated differently by the two methods: for IM3SHAPE

we use the reconstructed PSF image directly, and for NGMIX we fit models to the PSF.

A direct galaxy-by-galaxy test is not appropriate for a cross-catalog comparison, since there is not a unique unbiased shear estimate for a single galaxy. Rather, we wish to test that both methods produce consistent shear statistics for an ensemble of galaxies (cf. Velander et al. 2011). Two potential shear statistics that can be used are a galaxy-galaxy lensing signal and the two-point shear correlation function. We test if the results are consistent when using the same ensemble of galaxies with the same weighting.

Disagreement between the catalogs would be proof that at least one catalog is biased, but we would not be able to determine which one, nor the magnitude of this bias. Agreement between the two catalogs is subjectively reassuring, but we wish to emphasize that agreement does not prove that both catalogs are "correct" in the sense that they can be used to generate unbiased shear estimates.

8.6.1. Tangential shear ratio

Galaxy-galaxy lensing provides one of the cleanest tests of the relative calibration of the two catalogs, because the azimuthal symmetry inherent in the tangential shear signal largely cancels most sources of additive systematic error. Thus the ratio of two tangential shear signals is primarily a measure of the relative multiplicative errors between the two catalogs.

For this test, we use the tangential shear signal around Luminous Red Galaxies (LRGs) as determined by redMaGiC (red sequence Matched-filter Galaxies Catalog; Rozo et al. 2015) from the same DES SPT-E data. For this purpose, we do not require sources to be behind the lenses. Rather, we take the full LRG catalog as the lenses, and for the sources, we use all galaxies that are well-measured by both NGMIX and IM3SHAPE. Regardless of the redshifts of the LRGs and the source galaxies, the signal is expected to be the same for both catalogs.

The observed signal $\langle e_{t,i}(\theta)\rangle$ for each method $i \in \{\text{IM3SHAPE, NGMIX}\}$ can be written as:

$$\langle e_{t,i}(\theta)\rangle = (1+m_i)\langle\gamma_t(\theta)\rangle + \langle\eta_i(\theta)\rangle, \qquad (8\text{-}5)$$

where $\langle\gamma_t\rangle$ is the true underlying signal, $\langle\eta_i\rangle$ is a noise term including both intrinsic shape noise and measurement noise, and m_i is a possible calibration error for each method. We mostly drop the argument θ in the following for brevity. For the same ensemble of galaxies, the two catalogs have identical values

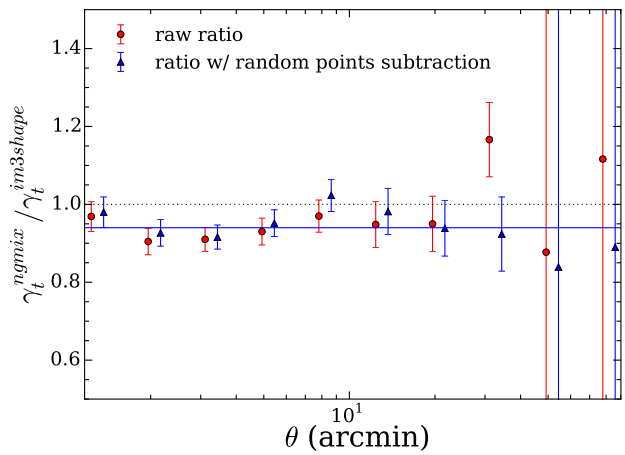

Fig. 25.— The ratios of tangential shear measurements around LRG galaxies from shears measured by NGMIX to those measured by IM3SHAPE. The red circles show the direct ratio and the triangles correspond to the ratio after subtraction of the tangential shear around random points. The weighted mean ratio in the scale range $1-20$ arcminutes is 0.954 ± 0.018. The blue line shows a prediction of the ratio (0.94) based on the GREAT-DES simulation, which accounts for a selection bias induced by the intersection of the two shape catalogs. This result is in good agreement with the data points.

of $\langle\gamma_t\rangle$ and a similar shape noise contribution to $\langle\eta_i\rangle$ (though not identical, since the two methods use different bands). The contribution to $\langle\eta_i\rangle$ from shape measurement noise, however, is expected to be somewhat different.

The red points in Figure 25 represent the ratio of measured tangential shear using the two shear catalogs. The weighted mean of the ratio over the range from 1 to 20 arcminutes (the typical scales of interest for weak lensing) is 0.932 ± 0.018. We would naively expect this to be an estimate of $(1+m_{\text{NGMIX}})/(1+m_{\text{IM3SHAPE}}) \approx 1 + m_{\text{NGMIX}} - m_{\text{IM3SHAPE}}$. However, three corrections are required before any conclusions can be drawn from this result about potential differences in the relative calibration.

First, additive systematic errors only cancel if the sources are distributed uniformly around the lenses. This is approximately true, but masking can break the symmetry, especially at large scales. One solution is to subtract off the measured tangential shear around random points, drawn from the same region and with the same masking as the LRGs. No signal is expected around such points, but any additive bias will affect both measurements equally. Thus the difference is a

cleaner estimate of the true tangential shear than the uncorrected signal. The blue points in Figure 25 show the effect of this subtraction, and have a mean ratio of 0.954 ± 0.018,

Second, the ratio of two noisy quantities with the same mean does not in general have an expectation value equal to 1. If the denominator is a random variable, X, with a symmetric probability distribution (e.g. $X \sim \mathcal{N}(\bar{X}, \sigma_X)$), the ratio will be approximately $1 + \sigma_X^2/\bar{X}^2$. To account for this bias, we create simulated realizations of the ratio, and compare the measured signal to the mean and variance of these. We generate a ratio realization in the following way:

1. Fit a polynomial, $\log(\langle e_t \rangle(\theta)) = p(\log(\theta))$ to the measured NGMIX signal, and take this to be the true signal, $\hat{\gamma}_t(\theta)$.

2. For each source in the ensemble, rotate both the NGMIX and IM3SHAPE shear by the same random angle.

3. Re-measure the two tangential shear signals, which now give estimates of the noise, $\langle \eta^r(\theta) \rangle$, as the true signal is removed by the random rotations.

4. Compute the realization ratio as

$$(\hat{\gamma}_t + \langle \eta^r_{\rm NGMIX} \rangle)/(\hat{\gamma}_t + \langle \eta^r_{\rm IM3SHAPE} \rangle). \qquad (8\text{-}6)$$

We find the mean of these realizations to be consistent with a ratio of 1 on all scales, and so conclude that the high $S/N$of the tangential shear ensures that the effect of the noise term in the denominator is negligible.

Finally, we find that the act of matching the two catalogs causes a selection bias in the NGMIX catalog for two reasons. First, the IM3SHAPE algorithm tends to fail more often for objects with low Sérsic index ($n < 1$). And second, the cuts we make on the IM3SHAPE measurements of $(S/N)_w$ and R_{gp}/R_p also affect the mix of galaxy properties in the matched catalog. These two selection effects, when applied to the NGMIX catalog impart a net bias on the NGMIX shear estimates in the matched catalog that is not present in the full NGMIX catalog.

We quantify the level of this selection bias by performing the same procedure on the GREAT-DES simulation. We compare the mean bias for NGMIX as a function of redshift first using only its own selection criteria and then also applying the IM3SHAPE selection. The result is shown in Figure 26. The matching induces a mean selection bias of about -3%. Furthermore, the bias increases with redshift. Weighting the bias according to the lens redshift distribution and the lensing efficiency of the source galaxies used in the tangential shear ratio test (and assuming that the lenses do not evolve with redshift), we find a net selection bias of -6% for NGMIX in the matched catalog relative to whatever bias might be present in the full NGMIX catalog[19].

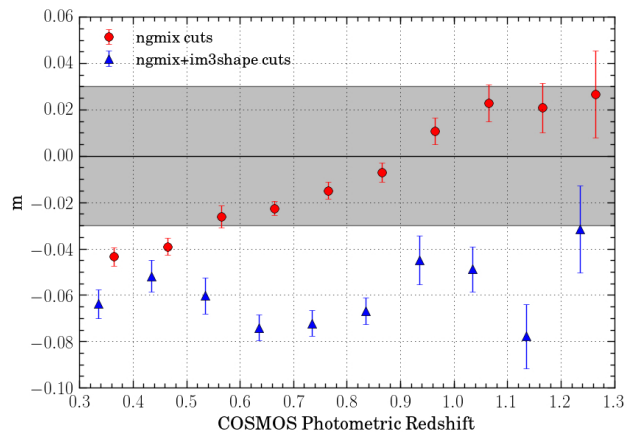

Fig. 26.— Multiplicative bias for NGMIX shear measurements on GREAT-DES simulated data as a function of redshift. The red circles show the bias calculated using all galaxies that pass the NGMIX selection criteria (as in the upper panel of in Figure 24). The blue triangles show the bias when also including the recommended IM3SHAPE selection, as we do to obtain the matched catalog used for Figure 25. As in Figure 24, the grey band represents the $\pm 3\%$ requirement for the SV data.

The mean ratio of 0.954 ± 0.018 is thus consistent with the prediction from GREAT-DES of -6% selection bias (which would produce a ratio of 0.94). This bias induced by the combination of IM3SHAPE and NGMIX selection criteria in the matched shape catalogs is shown by the blue line in Figure 25. Our finding is therefore consistent with no multiplicative bias in either catalog.

We cannot of course prove that neither catalog is affected by a significant multiplicative bias based on this test. They could both be biased by the same amount in either direction. Furthermore, there are significant un-

[19] We tested for a similar selection bias in the IM3SHAPE catalog due to imposition of the NGMIX cuts. The impact of the matching was found to be negligible, in part because the NGMIX catalog is deeper, so its cuts have very little impact on the IM3SHAPE selection.

certainties in the calculation of the predicted selection bias described above that may be at the $\sim 3\%$ level.

8.6.2. *Differential shear correlations*

The two-point shear correlation function is much more sensitive to additive shear errors than the tangential shear, as mentioned above; it would be difficult to disentangle multiplicative and additive errors in a ratio test. Even in the absence of additive errors, the ratio of shear correlation functions is much noisier than the ratio of tangential shears, making it a less stringent test of calibration.

For these reasons, we instead use the two point function of the *difference* in the shear estimates from NGMIX and IM3SHAPE to compare the shear catalogs:

$$\xi_{+,\Delta e}(\theta) = \langle (e_{\rm NGMIX}(\mathbf{x}) - e_{\rm IM3SHAPE}(\mathbf{x}))^* \\ (e_{\rm NGMIX}(\mathbf{x}+\boldsymbol{\theta}) - e_{\rm IM3SHAPE}(\mathbf{x}+\boldsymbol{\theta}))\rangle. \quad (8\text{-}7)$$

Consider the following model for the additive systematic errors in each catalog (labeled i here):

$$e_i = (1+m_i)\gamma + \eta_i + a_i c_{\rm common} + c_i, \quad (8\text{-}8)$$

where m_i is the calibration error, η_i is the noise in the estimate, $c_{\rm common}$ includes any additive systematic errors present in both catalogs, possibly multiplied by different coefficients a_i, and c_i is the additive error particular to each catalog.

By construction, the additive bias terms in equation 8-8 are independent. If we further make the assumption that the systematic errors are uncorrelated with the applied shear and the noise, and that m and c are uncorrelated, we find that

$$\begin{aligned}\xi_{+,\Delta e}(\theta) = {} & (\Delta m)^2 \xi_+(\theta) \\ & + (\Delta a)^2 \langle c^*_{\rm common} c_{\rm common}\rangle(\theta) \\ & + \langle c^*_{\rm NGMIX} c_{\rm NGMIX}\rangle(\theta) \\ & + \langle c^*_{\rm IM3SHAPE} c_{\rm IM3SHAPE}\,\rangle(\theta). \end{aligned} \quad (8\text{-}9)$$

This test is sensitive to the spatial correlations of the systematic errors in either catalog, but particularly to additive errors, rather than multiplicative. The $(\Delta m)^2$ factor for the multiplicative term typically makes this term insignificant.

There is one subtlety in the construction of this test. As we found in §8.6.1, the act of matching the two catalogs can induce selection biases that are not present in either catalog separately when using its own individual selection criteria. In this case, the salient selection effects are a spurious PSF leakage α and an overall mean $\langle c\rangle$ that can be induced by the match.

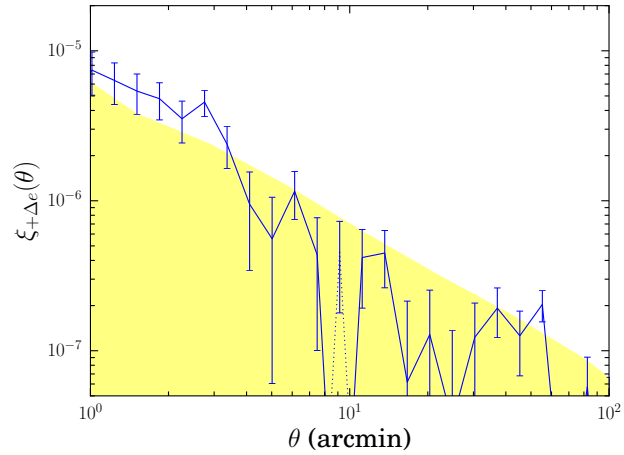

Fig. 27.— The shear auto-correlation function of the difference in shear estimates of NGMIX and IM3SHAPE. This test shows the level of additive systematic errors that may still be present in one catalog that is not present in the other. The yellow band is the requirement, $\delta\xi_+^{\rm max}$ from Figure 3.

The estimated value of α for NGMIX changes by less than 0.1% on the matched catalog relative to the full NGMIX catalog. But for IM3SHAPE, the matching changes α by -1.5%. Therefore, to make this a fair test of the additive systematic errors, we add back $0.015 \times e_{\rm PSF}$ to the IM3SHAPE galaxy shapes to account for this selection effect.[20]

Even after correcting for this, we also find that the mean shear changes by $(3.9+2.2i)\times 10^{-4}$ for NGMIX and by $(2.0-3.0i)\times 10^{-4}$ for IM3SHAPE. We ascribe these changes in the mean to be due to selection biases from the matching itself, leading to a spurious overall $\langle c\rangle$ for each catalog. We thus subtract these values as well from the shape estimates in each catalog.

Figure 27 shows the resulting correlation function (equation 8-7) after subtracting these selection biases. For the weights, we use $w = \sqrt{w_{\rm NGMIX} \times w_{\rm IM3SHAPE}}$. The yellow band is our requirement for additive systematic errors from equation 3-12. We see that at scales less than 3 arcminutes we are not quite meeting the requirements. Either one or both catalogs apparently have non-negligible additive systematic errors at these scales. We recommend that science applications

[20] We also subtract the corresponding value for NGMIX, although it makes no discernible difference.

Test	Upper Limit on Systematic Error	
	IM3SHAPE	NGMIX
PSF Model Tests		
§4.2 Mean PSF size error	$\|m\| < 0.005$	$\|m\| < 0.01$
§4.4 PSF model diagnostics $\rho_{1,3,4}$	$\xi_+^{cc}(1') < 2 \times 10^{-6\dagger}$	$\xi_+^{cc}(1') < 5 \times 10^{-6}$
	$\xi_+^{cc}(30') < 7 \times 10^{-8}$	$\xi_+^{cc}(30') < 9 \times 10^{-8}$
§4.4 PSF model diagnostics $\rho_{2,5}$	$\xi_+^{cc}(1') < 2 \times 10^{-7}$	$\xi_+^{cc}(1') < 8 \times 10^{-8}$
	$\xi_+^{cc}(30') < 1.5 \times 10^{-7}$	$\xi_+^{cc}(30') < 1.4 \times 10^{-7}$
Spatial Tests		
§8.1.1 Position in the field of view	No evidence of systematic errors	No evidence of systematic errors
§8.1.2 Position on CCD	$\xi_+^{cc}(1') < 6 \times 10^{-7}$	$\xi_+^{cc}(1') < 6 \times 10^{-7}$
§8.1.3 Tangential shear around field centers	$\xi_+^{cc}(1') < 4 \times 10^{-8}$	$\xi_+^{cc}(1') < 4 \times 10^{-8}$
PSF Tests		
§8.2.1 PSF leakage	$\|\alpha\| < 0.04$	$\|\alpha\| < 0.01$
§8.2.1 Dependence on PSF size	No evidence of systematic errors	No evidence of systematic errors
§8.2.2 Star-galaxy cross correlation	$\|\alpha\| < 0.03$	$\|\alpha\| < 0.015$
§8.2.3 Tangential shear around stars	No evidence of systematic errors	No evidence of systematic errors
Galaxy Property Tests		
§8.3.1 Galaxy S/N	No evidence of systematic errors	No evidence of systematic errors
§8.3.2 Galaxy size	No evidence of systematic errors	No evidence of systematic errors
B-mode Statistics		
§8.4 $\ell^2 C_{\rm BB}(\ell)/2\pi$	No evidence of B-mode	No evidence of B-mode
Calibration Tests		
§8.5 Redshift dependence in GREAT-DES	$\|m\| < 0.02$	$\|m\| < 0.04$
Cross-catalog Comparison		
§8.6.1 Tangential shear ratio	$\|\Delta m\| \lesssim 0.04$	
§8.6.2 Differential shear correlations	$\|\xi_+^{cc}(1')\| < 9 \times 10^{-6}$	
	$\|\xi_+^{cc}(30')\| < 2 \times 10^{-7}$	

† Since IM3SHAPE use only the r-band images, the values quoted here are based on the ρ statistics measured for the r-band-only PSFs. These curves are a bit higher than what is shown in Figure 9, which uses r, i, z bands.

Table 2: Summary of the results of our suite of null tests (including tests in §4.4). For reference, our nominal requirements from §3 are $|\alpha| < 0.03$, $|m| < 0.03$, $\xi_+^{cc}(1') < 7 \times 10^{-6}$, and $\xi_+^{cc}(30') < 2.5 \times 10^{-7}$.

sensitive to additive systematic errors be careful about checking how these small-scale systematic errors may affect their science results.

8.7. Summary of Systematics Tests

We now attempt to synthesize the results of our large suite of null tests. With this many tests, even if all the tests pass individually, it would not necessarily imply that the total systematic error is below our requirements. In this section we attempt to quantify an upper limit on the level of systematic error that may be in the shear catalogs, given all of the information we have available.

Table 2 provides a summary of the results from the previous sections (including the tests in §4). For each, we have converted the result of the test into the impact that the result could have on 4 possible values. For PSF leakage, we give the maximum allowed value of α. For other kinds of additive systematic errors, we give the maximum value of $\xi_+^{cc}(\theta)$ at $\theta = 1$ arcminute and (when relevant) 30 arcminutes. And for multiplicative errors, we give the maximum $|m|$ that is consistent with the test. Some tests do not lend themselves to a quantitative upper limit. Fortunately, in each of these cases, there is no evidence from the test that there is any systematic error.

There are two tests that give constraints on the PSF leakage coefficient α. In all cases, the tests are completely consistent with $\alpha = 0$. However, given the uncertainties in each case, we think it is appropriate to take the upper limit from the star-galaxy correlation function estimate, since it is the more precise estimate

in both cases. This gives us limits of $\alpha < 0.03$ for IM3SHAPE and $\alpha < 0.015$ for NGMIX. We can multiply this by ξ_+^{pp} to give a limit on the maximum additive systematic error we may have at $1'$ and $30'$ due to PSF leakage.

For the other additive systematic errors, we can add them together linearly. ξ acts like a variance, so systematic uncertainties add linearly, not in quadrature. However, the differential shear correlation test is different from the others. It includes many of the additive systematic errors tested by other tests, and in particular would almost certainly incorporate any systematic error due to PSF leakage, as the mechanism for any such leakage would be different for the two algorithms. Thus, it actually places a tighter limit on the potential systematic error from PSF leakage at $30'$ than the direct estimate of α.

The differential shear correlation does not however include all of the additive errors from the PSF model tests. The two codes use the same PSF model for the r-band exposures, although NGMIX also uses i and z-bands. We conservatively assume that the PSF modeling systematic errors act as $c_{\rm common}$ terms in the nomenclature of §8.6.2 and add them to the estimate from the differential shear correlation to get our final estimate on the possible additive systematic error in each catalog:

$$\text{IM3SHAPE} \quad \begin{array}{l} \left|\xi_+^{\rm sys}(1')\right| < 1.1 \times 10^{-5} \\ \left|\xi_+^{\rm sys}(30')\right| < 4 \times 10^{-7} \end{array} \tag{8-10}$$

$$\text{NGMIX} \quad \begin{array}{l} \left|\xi_+^{\rm sys}(1')\right| < 1.4 \times 10^{-5} \\ \left|\xi_+^{\rm sys}(30')\right| < 4 \times 10^{-7}. \end{array} \tag{8-11}$$

Note that we are not claiming that either catalog has systematic errors as large as this. Rather, we are claiming at $\sim 1\sigma$ level of confidence that the additive systematic errors in the two catalogs are smaller than this.

The limits on the multiplicative systematic errors come from two sources. We have estimated the bias on simulated data, and we have measured the relative bias of the two catalogs with respect to each other. With the exception of the lowest redshift bin for NGMIX, where we found a bias of $m \simeq -0.04$, all of the tests are consistent with $|m| < 0.02$ for both catalogs.

Investigation of the low redshift result for NGMIX indicates that it is largely due to that bin having more bulge galaxies than the higher-redshift bins, leading to increased model bias there. However, Figure 11 shows that the distribution of bulges in GREAT-DES may not match the data very well, in particular as a function of S/N, which is correlated with redshift. This makes us uncertain how applicable the $m = -0.04$ result is to the SV data.

Furthermore, while the tangential shear ratio test showed that the two catalogs were consistent to within $|\Delta m| < 0.02$, this was only after correcting for a matching-induced selection effect of $\Delta m \simeq 0.06$. This correction involves a number of assumptions, so we are not confident that it is more precise than about ± 0.03.

For these reasons, we feel that an appropriate upper limit on m for both catalogs is

$$|m| < 0.05. \tag{8-12}$$

We recommend science applications that are sensitive to multiplicative bias marginalize over a Bayesian prior on m centered at 0 with a standard deviation of 0.05.

9. Shear Catalogs

We plan to release the final shear catalogs publicly by January, 2016. See the Dark Energy Survey website[21] for instructions on how to download the catalogs.

In this section we describe the final galaxy selection, how to correctly apply the calibrations and sensitivities to ensembles of galaxies, and what final number density we achieve. Appendix C has further details about the content and structure of the catalogs.

9.1. Final Galaxy Selection

The starting point for our galaxy catalogs was described in §2.1 and §2.2. The former described how we selected regions of the survey where we trust the images, and the latter described our initial galaxy selection function. We now make further cuts informed by the suite of null tests in §8 such that the final shear catalog is found to pass these tests.

We remove individual objects according to the following criteria:

- SEXTRACTOR flags = 1 or 2. Objects with higher SEXTRACTOR flags have already been removed from the input catalogs, since they are clearly bad. But these two flags indicate that the object is likely to be blended, and thus the shape measurement is likely to be corrupted.
- "Crazy colors"[22]. Individual objects with questionable colors are probably contaminated by

[21] http://www.darkenergysurvey.org/

[22] "Crazy colors" mean any of the following: $g - r < -1$, $g - r > 4$, $i - z < -1$, or $i - z > 4$

cosmic rays or other defects, so their shapes are also likely to be bad.

- Very low surface brightness. We found a class of objects with very large sizes, but relatively low flux that were usually due to various image artifacts rather than real objects. We exclude objects with $i + 3.5\log(f_i/T) < 28$, where f_i is the i-band flux, and $T = I_{xx} + I_{yy}$ is the (deconvolved) object size estimated by NGMIX.

- Tiny size. If the NGMIX estimate of the object size is very small, then the object is probably a star. Specifically, we remove objects with $T + \sigma_T < 0.02$ square arcseconds.

From the resulting set of "good galaxies", we then make a further selection based on both S/N and the size of the galaxy relative to the PSF size, such that the resulting ensembles of shear estimates pass the null tests.

As we have already mentioned in §7.3.2, the IM3SHAPE selection needs to be made using $(S/N)_w$ and R_{gp}/R_p, since these are the parameters used for the shear calibration. NGMIX does not do any calibration, so its selection is made using $(S/N)_r$ and $T_r/T_{\rm PSF}$ (cf. equations 7-4 & 8-4) to avoid inducing a selection bias.

The selection that we find passes the suite of null tests is the following:

$$\begin{aligned} \text{IM3SHAPE}: \quad & (S/N)_w > 15 \\ & R_{gp}/R_p > 1.2 \\ \text{NGMIX}: \quad & (S/N)_r > 15 \\ & T_r/T_{\rm PSF} > 0.15. \end{aligned}$$

All of the test results shown in §8 use this selection.

Figure 28 shows the effect that successively applying each round of selections has on the distribution of r-band magnitudes, starting with the original Gold Catalog, selecting possible galaxies, removing problematic galaxies, and then applying the S/N and size criteria for the two shear catalogs.

9.2. Applying the Calibration/Sensitivity

For both IM3SHAPE and NGMIX, the raw galaxy shape values given in the catalog are intrinsically biased estimators of the shear. In the case of IM3SHAPE, simulation-based calibration is used (cf. §7.3.2). For NGMIX, the expectation value of the ellipticity was estimated from the posterior likelihood surface with a centrally-concentrated prior applied, which reduces the sensitivity of the estimator to an applied shear. An estimate of this sensitivity is calculated and given in the catalog (cf. §7.4.3).

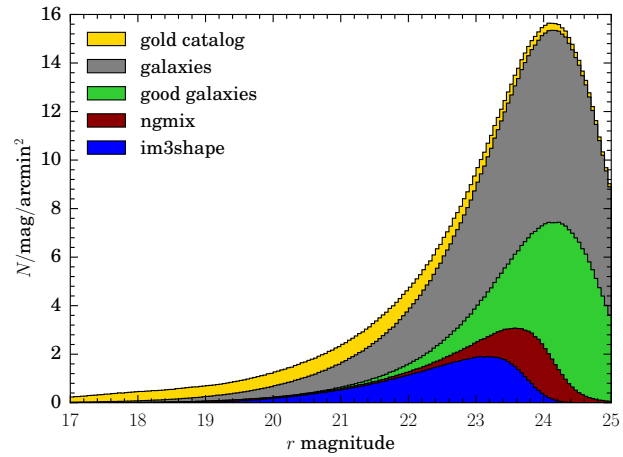

Fig. 28.— A histogram of the r-band magnitude distribution showing the application of the various selection criteria from the initial "Gold" Catalog to the two final shear catalogs. The dark red and blue show the galaxies with sufficiently accurate shape for NGMIX and IM3SHAPE respectively.

In both cases, the correction factor is a noisy estimate of the true correction. It is therefore not advisable to correct each galaxy's shape by the corresponding correction factor directly as this will introduce a bias. Rather the mean shear of an ensemble of galaxies should be corrected by the mean of the correction factors:

$$\langle\gamma\rangle = \frac{\sum(e_i - c_i)}{\sum s_i}, \tag{9-1}$$

where c_i is the additive correction for IM3SHAPE ($c_i \equiv 0$ for NGMIX) and s_i is the multiplicative correction $1 + m$ for IM3SHAPE or the estimated sensitivity for NGMIX.

The corrections in both cases are accurate in the limit of large numbers of galaxies. In practice, the ensemble should contain at least hundreds to thousands of galaxies to avoid dividing by noisy estimates of the mean sensitivity or shear bias correction.

In addition, each catalog comes with a recommended weight w_i to use for making these ensemble averages:

$$\langle\gamma\rangle = \frac{\sum w_i(e_i - c_i)}{\sum w_i s_i}. \tag{9-2}$$

For statistics such as tangential shear, you would apply the correction separately in each bin where you are

computing a mean shear. This will apply the appropriate correction to the subset of the galaxies that fall into each bin.

The correction method is slightly more complicated for two-point correlation functions, since each product involves two correction factors. In this case, the proper estimate is

$$\langle \gamma^a \gamma^b \rangle = \frac{\sum w_i^a w_j^b (e_i^a - c_i^a)(e_j^b - c_j^b)}{\sum w_i^a w_j^b s_i^a s_j^b}. \tag{9-3}$$

The denominator is just the two-point function of the scalar numbers s^a and s^b. The ratio is then taken for each bin in θ.

9.3. Effective Number Density

The effective number density of a weak lensing survey is defined implicitly in terms of the expected variance of either component of the estimated mean shear over its solid angle Ω (Chang et al. 2013):

$$\mathrm{var}(\langle \gamma_{1,2} \rangle) \equiv \frac{\sigma_{\mathrm{SN}}^2}{\Omega n_{\mathrm{eff}}}, \tag{9-4}$$

where σ_{SN} is the shape noise per component.

Applying all of the selections defined in §9.1 to our shape catalogs results in 2.12 million galaxies for IM3SHAPE and 3.44 million galaxies for NGMIX. The total useable area of SPTE is $\Omega = 139$ square degrees (cf. §2.1), which leads to direct number densities of 4.2 and 6.9 galaxies per square arcminute respectively.

To turn these numbers into proper effective number densities, we first need to calculate the shape noise σ_{SN}^2.

$$\sigma_{\mathrm{SN}}^2 = \frac{\sum w_i^2 \left(|e_i|^2 - 2\sigma_{e,i}^2 \right)}{2 \sum w_i^2 s_i^2}, \tag{9-5}$$

where $2\sigma_e^2$ is the trace of the covariance matrix of e_1, e_2[23], and the 2 in the denominator is to match the standard convention of quoting shape noise *per component*. As described above, s_i is the calibration factor or sensitivity correction for the two catalogs. For the IM3SHAPE catalog, this number comes to $\sigma_{\mathrm{SN}} = 0.233$, and for NGMIX, $\sigma_{\mathrm{SN}} = 0.243$.

The variance of each component of the mean shear over the entire survey area can be calculated from equation 9-2:

$$\mathrm{var}(\langle \gamma_{1,2} \rangle) = \frac{\sum w_i^2 \left(s_i^2 \sigma_{\mathrm{SN}}^2 + \sigma_{e,i}^2 \right)}{\left(\sum w_i s_i \right)^2}, \tag{9-6}$$

which, using equations 9-4 & 9-5, leads to

$$n_{\mathrm{eff}} = \frac{1}{\Omega} \frac{\sigma_{\mathrm{SN}}^2 \left(\sum w_i s_i \right)^2}{\sum w_i^2 \left(s_i^2 \sigma_{\mathrm{SN}}^2 + \sigma_{e,i}^2 \right)}. \tag{9-7}$$

$$= \frac{1}{\Omega} \frac{\left(\sum w_i s_i \right)^2}{\sum w_i^2 s_i^2} \left(1 - \frac{2 \sum w_i^2 \sigma_{e,i}^2}{\sum w_i^2 |e_i|^2} \right) \tag{9-8}$$

For IM3SHAPE, we find n_{eff} = 3.7 galaxies per square arcminute, and for NGMIX, n_{eff} = 5.7 galaxies per square arcminute.

Note that other authors use different definitions of n_{eff} than this. For instance, Heymans et al. (2012b) uses the definition

$$n_{\mathrm{eff}} = \frac{1}{\Omega} \frac{\left(\sum w_i \right)^2}{\sum w_i^2}. \tag{9-9}$$

Using this definition we obtain $n_{\mathrm{eff}} = 4.1$ and 6.8 for IM3SHAPE and NGMIX, respectively. With this definition however, the appropriate numerator in the ratio $\sigma_\epsilon^2 / n_{\mathrm{eff}}$ is not the intrinsic shape noise σ_{SN}^2, but rather the total shear noise including measurement noise. For our data, the values to use would be $\sigma_\epsilon = 0.245$ for IM3SHAPE and $\sigma_\epsilon = 0.265$ for NGMIX.

These number densities are quite a bit below the 10 galaxies per square arcminute that was predicted for DES (The Dark Energy Survey Collaboration 2005). This is in part because of our decision to cut both catalogs at $S/N > 15$ rather than 10 as we had originally hoped to be able to do. This removed about 0.5 million galaxies from the IM3SHAPE catalog and 1.0 million from the NGMIX catalog. Moving the IM3SHAPE size cut down to $R_{gp}/R_p > 1.15$ as well would add another 0.8 million galaxies. We hope that algorithm improvements to both catalogs will make these looser selection criteria possible in future DES analyses.

Furthermore, the average depth of the SV survey was not the full ∼10 exposures we expect for DES after five years. Instead, the mean is approximately 7 exposures averaged across the SPT-E area. If we reach an average of 10 exposures, This will lead to a 20% increase in the mean S/N and a corresponding increase in the number of usable galaxies.

In addition, the predicted value was based on an expected median seeing of 0.9″, while the median seeing

[23] IM3SHAPE does not produce a useful estimate of the covariance matrix, so we instead estimate σ_e^2 from the weights, which are designed to be an estimate of $1/(\sigma_{\mathrm{SN}}^2 + \sigma_e^2)$ (cf. §7.3.4).

during science verification was slightly above $1.0''$. We are closer to achieving our goal of $0.9''$ in the main survey observations (Diehl et al. 2014), so this will help to increase $n_{\rm eff}$.

Another reason for the low number count is the rejection of objects with neighbors. The SEXTRACTOR flags related to blended objects removed almost 1 million galaxies from the catalogs. We are currently working on an algorithm to model the profiles of neighboring objects so their light profiles can be effectively removed from the image and not contaminate the shapes of nearby objects, thus allowing us to keep more of these objects in the catalog.

Another obvious improvement will be to use multi-band fitting in IM3SHAPE, which would increase the S/N of each galaxy by using more pixels of information. This is already implemented, but it was not complete in time to be run and tested on these data. It will be used in the next DES analysis.

Finally, the detection of image artifacts in the data management pipeline has been improved from the version used for the SV data. The removal of low surface brightness objects, which was designed to remove a large proportion of these artifacts, removed 1.5 million objects. Presumably many of these are real galaxies rather than image artifacts, so if we can omit this step, we will keep more galaxies in the catalog.

With all of these improvements to both the data and the algorithms, we are optimistic that we will be able achieve our forecasted $n_{\rm eff} = 10$ galaxies per square arcminute in the 5-year DES analysis.

10. Summary and Discussion

We present here two shear catalogs for the SPT-E region observed as part of the DES science verification time. Both catalogs, NGMIX and IM3SHAPE, have passed a comprehensive suite of null tests that show that they are accurate enough to be used for weak lensing science with these data. The catalogs have 4.2 and 6.9 galaxies per square arcminute (for IM3SHAPE and NGMIX respectively), which corresponds to 2.12 and 3.44 million galaxies over the 139 square degree footprint. These correspond to effective number densities of 3.7 and 5.7 galaxies per square arcminute, respectively (cf. §9.3).

Both shear catalogs use the original single-epoch pixel data to jointly constrain the galaxy models, thereby avoiding issues of correlated noise and complex PSF interpolation that occur when using stacked images. This is a relatively new technique in weak lensing, having only previously been employed on real data by Heymans et al. (2012b) and Kuijken et al. (2015). However, given its significant advantages, we believe it will become the standard algorithm employed by most future surveys.

In addition to passing individual null tests on the data, the two catalogs are shown to be consistent with each other, both in terms of possible additive systematic errors and the overall calibration (i.e. multiplicative systematic errors). This is a non-trivial result, considering that the calibration strategies of the two catalogs are completely different; IM3SHAPE calibrates the shear bias from simulations, and NGMIX uses a Bayesian algorithm that is relatively insensitive to noise bias, but does require a prior on the ellipticity distribution. This is the first significant weak lensing analysis to present two accurate and independent shear catalogs, and thus the first to be able to show this kind of consistency.

In §8.7 we estimate upper limits on the level of additive systematic errors that may be present in the two catalogs at $1'$ and $30'$. We also recommend a Bayesian prior of $|m| < 0.05$ for the systematic uncertainty on the calibration for both catalogs.

While the catalogs are seen to be sufficiently accurate for SV weak lensing science, they are not yet passing the tests at the level that will be needed for the full 5-year DES data. There is still a significant amount of work required to improve the algorithms to meet those requirements.

One area that needs improvement is our PSF modeling (cf. Figures 6 & 9). Fortunately, there has been a significant amount of work in recent years on improved PSF modeling and interpolation algorithms (e.g. Chang et al. 2012; Li et al. 2012; Kitching et al. 2013; Gentile et al. 2013). We have also been working on an algorithm to model the PSF using the actual optical aberrations measured from the wavefront sensors in the corners of the DECam field of view (Roodman et al. 2014). We will investigate whether incorporating this information can lead to more accurate PSF interpolation.

We also expect a significant improvement in the astrometric solution in the next round of analysis. It will include a more accurate functional form for the telescope distortion and also take into account effects like edge distortions and tree rings that are present in the data (cf. Figure 15). We expect this to reduce some of the spurious features seen in Figures 8 & 16.

We have recently implemented an algorithmic correction to the brighter-fatter relation discussed in §4.2 (Gruen et al. 2015). This will allow us to use brighter stars for constraining the PSF than we were able to use in this analysis, which is expected to lead to better estimated PSFs.

We are working on an improved algorithm for handling neighbors by subtracting off an estimate of their light profile rather than merely masking contaminated pixels. While not a perfect subtraction, we expect this will let us use more pixels for constraining the galaxy models, which will lead to fewer galaxies being removed from the final catalogs. Contamination by neighbors was one of the more significant cuts that led to the drop in number density for the "good galaxies" seen in Figure 28.

There are also two new shear algorithms being developed for DES. One is based on the Bayesian Fourier domain (BFD) algorithm of Bernstein & Armstrong (2014). The other is based on the MetaCalibration strategy presented in Mandelbaum et al. (2015). Considering how useful we found it to have two catalogs, we are looking forward to the prospect of additional catalogs to compare in various ways.

We also plan to start implementing corrections for the chromatic effects of the PSF described by Meyers & Burchat (2015). According to their estimates of the effects of PSF chromaticity, it is not expected to be a significant problem for the current analysis, but we will need to correct for these effects in the 5-year data analysis.

In addition to these planned algorithmic improvements, the data itself will be somewhat better in the main survey. Part of the reason for taking the science verification data was to find problems with the camera and telescope hardware. As such, quite a few hardware improvements were made during this time, as well as some in the following year (cf. Diehl et al. 2014). The image quality for the main survey is thus significantly better than the already quite good image quality in the SV data.

We therefore believe that we will be able to significantly improve the quality of the shear catalogs in future DES analyses. We need to keep the level of systematic errors below the improved statistical uncertainty for these data. The full 5-year DES data will cover about 30 times more area, so our requirements for the systematic errors will drop by roughly a factor of 5. By implementing the improvements discussed here, we hope to keep pace with the requirements.

Acknowledgements

We are grateful for the extraordinary contributions of our CTIO colleagues and the DECam Construction, Commissioning and Science Verification teams in achieving the excellent instrument and telescope conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the DES Data Management group.

We thank the many DES-internal reviewers whose suggestions have vastly improved this paper during the collaboration-wide review process. The DES publication number for this article is DES-2015-0059. The Fermilab preprint number is FERMILAB-PUB-15-309-AE.

Jarvis has been supported on this project by NSF grants AST-0812790 and AST-1138729. Jarvis, Bernstein, Clampitt, and Jain are partially supported by DoE grant DE-SC0007901. Sheldon is supported by DoE grant DE-AC02-98CH10886. Zuntz, Kacprzak, Bridle, and Troxel acknowledge support from the European Research Council in the form of a Starting Grant with number 240672. Das was funded by DoE Grant DE-SC0007859. Gruen was supported by SFB-Transregio 33 'The Dark Universe' by the Deutsche Forschungsgemeinschaft (DFG) and the DFG cluster of excellence 'Origin and Structure of the Universe'. Gangkofner acknowledges the support by the DFG Cluster of Excellence 'Origin and Structure of the Universe'. Melchior was supported by DoE grant DE-FG02-91ER40690. Plazas was supported by DoE grant DE-AC02-98CH10886 and by JPL, run by Caltech under a contract for NASA. Lima is partially supported by FAPESP and CNPq.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the

Ministério da Ciência e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The DES data management system is supported by the National Science Foundation under Grant Number AST-1138766. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986, and Centro de Excelencia Severo Ochoa SEV-2012-0234, some of which include ERDF funds from the European Union.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Edinburgh, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

A. Multi-Epoch Data Structures

A.1. Bulding the MEDS Files

Each MEDS file corresponds to a single coadd image. For each one, we gather the list of all single-epoch images that were used to construct the coadd image. Then for each object in the corresponding coadd detection catalog, we identify the location of the object in all single-epoch images where that object appears using each image's WCS transformation to convert between the coordinate systems. We then identify a region around each object in each single-epoch image and save it as a postage stamp in the MEDS file. A postage stamp from the coadd image is also stored in the file as the first entry for each object.

The size of the cutout is determined from the basic SEXTRACTOR measurements `FLUX_RADIUS`, `A_WORLD` and `B_WORLD` as follows:

$$s = 2 \times 5 \times \sigma \times (1 + \epsilon) \tag{A1}$$

$$\sigma = \mathrm{FWHM}/fac \tag{A2}$$

$$\mathrm{FWHM} = 2 \times \texttt{FLUX_RADIUS} \tag{A3}$$

$$\epsilon = 1 - \texttt{B_WORLD}/\texttt{A_WORLD}, \tag{A4}$$

where $fac \sim 2.35$ is the conversion between FWHM and σ. The `FLUX_RADIUS` is a robustly measured quantity, being the radius of the circular aperture enclosing half the estimated total flux of the object. We find that `A_WORLD` and `B_WORLD`, while not suitable for a lensing shear analysis, are measured well enough for the purpose of estimating the eccentricity ϵ.

We take the maximum of the size s from all single-epoch measurements as the fiducial cutout size. To facilitate fast FFT calculations on the cutouts, we round the fiducial cutout sizes upward to either a power of two or 3 times a power of two.

In addition to the image cutouts, we also store the SEXTRACTOR weight map and segmentation map. The cutouts are background subtracted using the background maps output by SEXTRACTOR. The weight maps are modified to be zero anywhere that a flag is set in the SEXTRACTOR maskplane, which includes defects such as bad pixels. The different image types are stored in separate extensions of the file, along with a plethora of metadata. See §A.2 for details.

All images, including the coadd image, are placed on the same photometric system such that the magnitude zero point is 30.0. The weight maps are also scaled appropriately.

Because the full set of data to be stored in the MEDS file does not fit into memory simultaneously, we use the ability of CFITSIO[24] to write chunks of images directly to disk without keeping the full image in main memory.

The code for creating MEDS files, including the WCS transformation library, is hosted publicly as part of a larger package `deswl_shapelets`[25]. The code that generates the input object list, including cutout sizes, is part of the `meds` software library (see §A.2 for more details).

A.2. Structure of the MEDS Files

The MEDS data, including all of the images of each object observed in a single coadd tile, along with appropriate catalog information, are stored in a single FITS file with a number of extensions. Table 3 gives an overview of the FITS file structure.

To simplify access to the data in the MEDS filess, we provide an Application Programmer's Interface (API) library, `meds`, which is available for download[26] and is free software. A full API is provided for the Python language. A smaller subset of the full functionality is available as a library for the C programming language. For complete documentation, we direct the reader to the `meds` repository URL.

[24] `http://heasarc.gsfc.nasa.gov/fitsio/fitsio.html`

[25] `https://github.com/rmjarvis/deswl_shapelets`

[26] `https://github.com/esheldon/meds`

Extension Name	FITS Extension Type	Comments
object_data	binary table	Information for each coadd object
image_info	binary table	Information for each image
metadata	binary table	MEDS configuration data
image_cutouts	image	Image cutouts for coadd and single-epoch images
weight_cutouts	image	Weight map cutouts
seg_cutouts	image	Segmentation map cutouts
bmask_cutouts	image	Bitmask map cutouts

Table 3: Overview of the MEDS FITS file structure.

Column Name	Data Type [type][Nbytes]	Comments
id	i8	The id column from the input file
number	i8	SExtractor NUMBER column from coadd catalog
ncutout	i8	number of cutouts for this object
box_size	i8	box size for each cutout
file_id	i8[NMAX]	zero-offset id into the image_info extension
start_row	i8[NMAX]	zero-offset, start of cutout in the image extensions.
orig_row	f8[NMAX]	zero-offset position in original image
orig_col	f8[NMAX]	zero-offset position in original image
orig_start_row	i8[NMAX]	zero-offset start corner in original image
orig_start_col	i8[NMAX]	zero-offset start corner in original image
cutout_row	f8[NMAX]	zero-offset position in cutout image
cutout_col	f8[NMAX]	zero-offset position in cutout image
dudrow	f8[NMAX]	Jacobian matrix of WCS transformation
dudcol	f8[NMAX]	Jacobian matrix of WCS transformation
dvdrow	f8[NMAX]	Jacobian matrix of WCS transformation
dvdcol	f8[NMAX]	Jacobian matrix of WCS transformation

Table 4: Column definitions for the object_data MEDS file extension. Column data types specify iN for integer and fN for floating point types, where N is the number of bytes. Array columns are additionally marked with [NMAX], where NMAX is the maximum number of cutouts for any object.

Information about each coadd object is stored in the object_data FITS binary table extensions (see Table 4) 4. The id column is an arbitrary integer input with the MEDS catalog file. For the DES MEDS files we set this column to the coadd_objects_id, a unique identifier from the DES database. The Jacobian matrix is calculated from the WCS found in the FITS header of each image file. This matrix can be used to work in "sky coordinates" as opposed to the pixel grid; this is important to ensure all measurements are performed in the same coordinate system.

Information about each image is stored in the image_info binary FITS table extension. In Table 5 we show the column definitions. The image_id column is an arbitrary integer input with the file list. For our work we set this column to the image_id column from the DES database, a unique identifier for each single-epoch image. The magzp column is the original zero point for the image, defined in the FITS header. All images are scaled to have a common zeropoint, which is also an input the the MEDS creation code, and is stored in the metadata extension (see below). This image scaling factor is stored in the scale column.

All configuration information, including that sent on the command line, is stored in the metadata FITS binary table extension. A wide variety of configuration data is stored, but a few of the more useful columns are given in Table 6. The magzp_ref column is the common zeropoint used for all images, as described above. This is an input parameter for the MEDS creation code; for our work we use a zeropoint of 30.0.

The image cutouts are stored in the image_cutouts FITS image extension. Because the cutouts are variable in size, storage as a two-dimensional array would be awkward. We instead use a one-dimensional array strorage. All

Column Name	Data Type [type][Nbytes]	Comments
image_id	i8	an id for the image
image_flags	i8	flag bitmask for the image
image_path	SXX	full path to source image
wcs_path	SXX	full path to wcs file
sky_path	SXX	full path to sky image
seg_path	SXX	full path to segmentation image
magzp	f4	magnitude zero point for this image
scale	f4	scale factor used to place on common system

Table 5: Column definitions for the image_info MEDS file extension, with table entries defined as in Table 4 except the SXX data types, which indicate a string size can vary for each file.

Column Name	Data Type [type][Nbytes]	Comments
magzp_ref	f4	the reference zero point for all images
DESDATA	SXX	root directory for DES data
cat_file	SXX	the file with the box sizes and other input
coadd_file	SXX	the coadd file
coadd_image_id	SXX	the id of the coadd
coadd_srclist	SXX	the list of single-epoch sources
coaddcat_file	SXX	the coadd catalog file
coaddseg_file	SXX	the file holding the segmentation map
min_boxsize	i4	the minimum cutout size
max_boxsize	i4	the maximum cutout size
medsconf	SXX	the meds configuration identifier

Table 6: Selected column definitions for the metadata MEDS file extension, with table entries defined as in Tables 4 & 5.

cutouts for an objects are stored sequentially. The starting row, number of cutouts, and cutout box size are stored in the object_info table for each object, making retrieval straightforward. The MEDS software library, described above, makes retrieval transparent. The data are stored as single precision floating point.

The weight_cutouts, seg_cutouts and bmask_cutouts extensions are laid out exactly as the image_cutouts extension, and store the weight maps, segmentation maps and bitmask plane cutouts, respectively.

Finally, we note that the data compress very well, especially the segmentation and bitmask maps which respond well to run-length encoding. We use the fpack program[27], which reduces the file size by a factor of approximately five. The final files are about 200 MB per epoch for each MEDS file, each of which corresponds to one coadd tile.

B. Catalog Flags

B.1. IM3SHAPE Flags

IM3SHAPE reports two kinds of flags. Table 7 lists "error" flags, and Table 8 lists "info" flags. For the most conservative treatment, users can select galaxies where both are zero. However, using info flags > 0 may be appropriate in some cases.

[27] http://heasarc.gsfc.nasa.gov/fitsio/fpack/

Position	Value	Meaning
0	1	IM3SHAPE failed completely
1	2	Minimizer failed to converge
2	4	Tiny ellipticty $e < 1e-4$: IM3SHAPE fit failed
3	8	e1 or e2 outside $(-1, 1)$
4	16	Radius > 20 arcseconds
5	32	$R_{gp}/R_p > 6$ - huge galaxy
6	64	Negative or nan R_{gp}/R_p
7	128	$S/N < 1$
8	256	χ^2 per effective pixel > 3
9	512	Normed residuals < -20 somewhere
10	1024	Normed residuals > 20 somewhere
11	2048	δu more than 10 arcseconds from nominal
12	4096	δv more than 10 arcseconds from nominal
13	8192	Failed to measure the FWHM of PSF or galaxy
14	16384	r-band SEXTRACTOR flag has 0x4 or above

Table 7: Error flags in the IM3SHAPE catalog. Objects with non-zero error flag should be removed from any science analysis.

Position	Value	Meaning
0	1	r-band SEXTRACTOR flagged with 0x1, indicating bright neigbours
1	2	r-band SEXTRACTOR flagged with 0x2, indicating blending
2	4	Mask fraction > 0.5
3	8	Model image < -0.01 somewhere
4	16	$R_{gp}/R_p < 1.15$
5	32	Radius > 5 arcseconds
6	64	Centroid more than 0.6 arcseconds from nominal
7	128	χ^2 per effective pixel > 1.25
8	256	$R_{gp}/R_p > 3.5$ (very large galaxy)
9	512	Normed residuals < -2 somewhere
10	1024	Normed residuals > 2 somewhere
11	2048	Declination below limit where we have good photometry
12	4096	$S/N > 10000$
13	8192	Radius > 10 arcseconds
14	16384	$S/N < 10$
15	32768	Model image < -0.05 somewhere
16	65536	χ^2 per effective pixel < 0.8
17	131072	More than 70% of fitted flux is in masked region
18	262144	Model completely negative
19	524288	χ^2 per effective pixel > 2
20	1048576	Very large PSF
21	2097152	Negative PSF FWHM
22	4194304	$R_{gp}/R_p > 3.5$
23	8388608	Centroid more than one arcsecond from nominal
24	16777216	Mask fraction > 0.75
25	33554432	One or more error flags is set

Table 8: Info flags in the IM3SHAPE catalog. Objects with non-zero info flag may be acceptable depending on the scientific application.

B.2. NGMIX Flags

The NGMIX catalog has an error flag that indicate when some kind of error occurred during the fitting procedure. Users should only use galaxies with `error_flag == 0`. The meanings of the various possible error flag values are given in Table 9.

Position	Value	Meaning
0	1	`NO_CUTOUTS` There were no cutouts for this object
1	2	`PSF_FIT_FAILURE` PSF fitting failed for all epochs
2	4	Not used
3	8	`GAL_FIT_FAILURE` Galaxy fitting failed
4	16	`BOX_SIZE_TOO_BIG` Box size was larger than 2048
5	32	Not used
6	64	`LOW_PSF_FLUX` The S/N of the PSF flux was lower than 4 in all bands
7	128	`UTTER_FAILURE` Utter failure of the fitting. For this release, the flag was set when no valid guess for the fitters could be generated.
30	1073741824	`NO_ATTEMPT` No attempt of a fit was made due to other flags.

Table 9: Error flag values in the NGMIX catalog. Objects with non-zero error flag should be removed from any science analysis.

C. Details of the Shear Catalogs

There are three files containing the final DES SV shear catalogs:

- `des_sv_wl_im3shape.fits` is the IM3SHAPE catalog. The relevant columns in this catalog are listed in Table 10.
- `des_sv_wl_ngmix.fits` is the NGMIX catalog. The relevant columns in this catalog are listed in Table 11
- `des_sv_wl_info.fits` has flags that can be used to select a set of galaxies with good shear estimates. It also has columns with information from the main coadd catalog, such as RA and Declination, for convenience in using these catalogs without having to join them to the main DES object catalog. Photometric redshift information is based on the SkyNet algorithm (Sánchez et al. 2014; Bonnett 2015; Bonnett et al. 2015).

Most users will want to select objects where `sva1_flags == 0`. This selects the objects that we are confident are actually galaxies, and not either stars or some kind of spurious artifact in the data. See Table 13 for the meaning of non-zero values of this flag. In addition we have two additional columns that indicate which galaxies fail the IM3SHAPE and NGMIX selection criteria. The `im3shape_flags` column is zero if

```
(error_flag==0) & (info_flag==0) &
(snr>15) & (rgpp_rp>1.2)
```

The `ngmix_flags` column is zero if

```
(flags==0) & (exp_flags==0) &
(round_flags==0) & (exp_round_flags==0) &
(0.4<exp_arate<0.6) &
(exp_e_sens_1>0.0) & (exp_e_sens_2>0.0)
(exp_s2n_r>15) & (exp_T_r/psfrec_T>0.15)
```

In each case, these select the galaxies which have been found to pass all of the null tests in §8. Users can thus select galaxies with good shear estimates by simply selecting `im3shape_flags==0` or `ngmix_flags==0` as appropriate.

Column	Meaning
coadd_objects_id	A unique id number of the object
e_1	The raw e_1 shape estimate
e_2	The raw e_2 shape estimate
nbc_m	The multiplicative bias correction
nbc_c_i	($i \in \{1,2\}$) The additive bias corrections
w	The recommended weight
snr_w	The estimated $(S/N)_w$
snr_r	The estimated $(S/N)_r$
error_flag	The error flag (cf. Table 7 in Appendix B)
info_flag	The info flag (cf. Table 8 in Appendix B)

Table 10: The most relevant columns in the IM3SHAPE catalog.

Column	Meaning
coadd_objects_id	A unique id number of the object
e_1	The raw e_1 shape estimate
e_2	The raw e_2 shape estimate
sens_avg	The sensitivity estimate
w	The recommended weight
snr_w	The estimated $(S/N)_w$
snr_r	The estimated $(S/N)_r$
e_cov_i_j	($i, j \in \{1,2\}$) The covariance matrix of the shape estimate
error_flag	The error flag (cf. Table 9 in Appendix B)

Table 11: The most relevant columns in the NGMIX catalog.

Column	Meaning
coadd_objects_id	A unique id number of the object
sva1_flags	A flag indicating problematic galaxies (cf. Table 13)
im3shape_flags	A flag that is 0 if this object's shape in the IM3SHAPE catalog is good to use; 1 if not.
ngmix_flags	A flag that is 0 if this object's shape in the NGMIX catalog is good to use; 1 if not.
ra	The right ascension of the object in degrees
dec	The declination of the object in degrees
mag_auto_g	The g-band magnitude
mag_auto_r	The r-band magnitude
mag_auto_i	The i-band magnitude
mag_auto_z	The z-band magnitude
mean_photoz	A point estimate of the photometric redshift
photoz_bin	The cosmological photometric redshift bin (0,1,2)

Table 12: The most relevant columns in the info catalog.

Position	Value	Meaning
0	1	i-band SEXTRACTOR flag has bit 0 set, indicating possible blend
1	2	i-band SEXTRACTOR flag has bit 1 set, indicating definite blend
2	4	Modest Classification calls this object a star (`bright_test` or `locus_test` from §2.2)
3	8	Modest Classification calls this object junk (`faint_psf_test` from §2.2)
4	16	In region with high density of objects with "crazy colors"
5	32	In region with lots of points with large centroid shifts between bandpasses
6	64	Near a 2MASS star
7	128	Large offset in g and i band windowed positions
8	256	Object did not get measured by NGMIX
9	512	Likely star according to NGMIX $T + \sigma_T < 0.02$ square arcseconds
10	1024	Very low surface brightness according to NGMIX measurements
11	2048	Object does not satisfy good measurement flags in NGMIX

Table 13: Values of the `sva1_flags` in the info catalog

REFERENCES

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2004, AJ, 128, 502

Amara, A., & Réfrégier, A. 2008, MNRAS, 391, 228

Antilogus, P., Astier, P., Doherty, P., Guyonnet, A., & Regnault, N. 2014, Journal of Instrumentation, 9, C3048

Avery, P. 1996, Combining Measurements with Correlated Errors, CBX 95–55

Becker, M. R. 2013, MNRAS, 435, 1547

Becker, M. R., & Rozo, E. 2014, arXiv:astro-ph/1412.3851, arXiv:1412.3851

Becker, M. R., Troxel, M. A., MacCrann, N., et al. 2015, ArXiv e-prints, arXiv:1507.05598

Bergé, J., Gamper, L., Réfrégier, A., & Amara, A. 2013, Astronomy and Computing, 1, 23

Bernstein, G. M. 2010, MNRAS, 406, 2793

Bernstein, G. M., & Armstrong, R. 2014, MNRAS, 438, 1880

Bernstein, G. M., & Gruen, D. 2014, PASP, 126, 287

Bernstein, G. M., & Jarvis, M. 2002, AJ, 123, 583

Bertin, E. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 112

Bertin, E. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 435

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Bonnett, C. 2015, MNRAS, 449, 1043

Bonnett, C., Hartley, M. A. T. W., Amara, A., et al. 2015, ArXiv e-prints, arXiv:1507.05909

Bosch, J. F. 2011, PhD thesis, University of California, Davis

Bouy, H., Bertin, E., Moraux, E., et al. 2013, A&A, 554, A101

Bridle, S., Shawe-Taylor, J., Amara, A., et al. 2009, Annals of Applied Statistics, 3, 6

Bridle, S., Balan, S. T., Bethge, M., et al. 2010, MNRAS, 405, 2044

Bruderer, C., Chang, C., Refregier, A., et al. 2015, ArXiv e-prints, arXiv:1504.02778

Carlstrom, J. E., Ade, P. A. R., Aird, K. A., et al. 2011, PASP, 123, 568

Chang, C., Marshall, P. J., Jernigan, J. G., et al. 2012, MNRAS, 427, 2572

Chang, C., Jarvis, M., Jain, B., et al. 2013, MNRAS, 434, 2121

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2001, ApJ, 559, 552

de Vaucouleurs, G. 1948, Annales d'Astrophysique, 11, 247

Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977, Journal of the Royal Statistical Society, 39, 1

Desai, S., Armstrong, R., Mohr, J. J., et al. 2012, ApJ, 757, 83

Diehl, H. T. 2012, in Physics Procedia, Vol. 37, Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011), 1332 – 1340, proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011)

Diehl, H. T., Abbott, T. M. C., Annis, J., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9149, 91490V

Duda, R. O., & Hart, P. E. 1972, Commun. ACM, 15, 11

Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, ArXiv e-prints, arXiv:1504.02900

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9

Gentile, M., Courbin, F., & Meylan, G. 2013, A&A, 549, A1

Goodman, J., & Weare, J. 2010, App. Math. Comp. Sci., 5, 65

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Gruen, D., Bernstein, G. M., Jarvis, M., et al. 2015, ArXiv e-prints, arXiv:1501.02802

Gruen, D., Brimioulle, F., Seitz, S., et al. 2013, MNRAS, 432, 1455

Guyonnet, A., Astier, P., Antilogus, P., Regnault, N., & Doherty, P. 2015, A&A, 575, A41

Heymans, C., Rowe, B., Hoekstra, H., et al. 2012a, MNRAS, 421, 381

Heymans, C., Brown, M. L., Barden, M., et al. 2005, MNRAS, 361, 160

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323

Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012b, MNRAS, 427, 146

Hirata, C., & Seljak, U. 2003, MNRAS, 343, 459

Hirata, C. M., Mandelbaum, R., Seljak, U., et al. 2004, MNRAS, 353, 529

Hoekstra, H. 2013, ArXiv e-prints, arXiv:1312.5981

Hoffleit, D., & Jaschek, C. . 1991, The Bright star catalogue

Hogg, D. W., & Lang, D. 2013, PASP, 125, 719

Honscheid, K., Elliott, A., Annis, J., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451, 845112

Hough, P. V. C. 1959, in Proceedings of the International Conference on High Energy Accelerators and Instrumentation, Vol. 1

Jarvis, M., Bernstein, G. M., Fischer, P., et al. 2003, AJ, 125, 1014

Jee, M. J., Tyson, J. A., Schneider, M. D., et al. 2013, ApJ, 765, 74

Kacprzak, T., Bridle, S., Rowe, B., et al. 2014, MNRAS, 441, 2528

Kacprzak, T., Zuntz, J., Rowe, B., et al. 2012, MNRAS, 427, 2711

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460

Kannawadi, A., Mandelbaum, R., & Lackner, C. 2015, MNRAS, 449, 3597

Kitching, T., Balan, S., Bernstein, G., et al. 2010, ArXiv e-prints, arXiv:1009.0779

Kitching, T. D., Balan, S. T., Bridle, S., et al. 2012, MNRAS, 423, 3163

Kitching, T. D., Rowe, B., Gill, M., et al. 2013, ApJS, 205, 12

Klein, J. R., & Roodman, A. 2005, Annual Review of Nuclear and Particle Science, 55, 141

Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196

Krause, E., & Hirata, C. M. 2010, A&A, 523, A28

Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, ArXiv e-prints, arXiv:1507.00738

Leauthaud, A., Massey, R., Kneib, J.-P., et al. 2007, ApJS, 172, 219

Leistedt, B., Peiris, H. V., Elsner, F., et al. 2015, ArXiv e-prints, arXiv:1507.05647

Levenberg, K. 1944, Quarterly of Applied Mathematics, 164

Lewis, A. 2009, MNRAS, 398, 471

Li, G., Xin, B., & Cui, W. 2012, ArXiv e-prints, arXiv:1203.0571

Lourakis, M. 2004, levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++, `http://www.ics.forth.gr/~lourakis/levmar/`

Mandelbaum, R., Hirata, C. M., Leauthaud, A., Massey, R. J., & Rhodes, J. 2012, MNRAS, 420, 1518

Mandelbaum, R., Hirata, C. M., Seljak, U., et al. 2005, MNRAS, 361, 1287

Mandelbaum, R., Rowe, B., Bosch, J., et al. 2014, ApJS, 212, 5

Mandelbaum, R., Rowe, B., Armstrong, R., et al. 2015, MNRAS, 450, 2963

Marquardt, D. 1963, SIAM Journal on Applied Mathematics, 431

Massey, R., & Refregier, A. 2005, MNRAS, 363, 197

Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13

Melchior, P., Böhnert, A., Lombardi, M., & Bartelmann, M. 2010, A&A, 510, A75

Melchior, P., & Viola, M. 2012, MNRAS, 424, 2757

Melchior, P., Viola, M., Schäfer, B. M., & Bartelmann, M. 2011, MNRAS, 412, 1552

Melchior, P., Suchyta, E., Huff, E., et al. 2015, MNRAS, 449, 2219

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087

Meyers, J. E., & Burchat, P. R. 2015, ApJ, 807, 182

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & van Waerbeke, L. 2007, MNRAS, 382, 315

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858

Mohr, J. J., Armstrong, R., Bertin, E., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451, 84510D

Nakajima, R., & Bernstein, G. 2007, AJ, 133, 1763

Nelder, J. A., & Mead, R. 1965, Computer Journal, 308

Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A., & Bridle, S. L. 2008, A&A, 484, 67

Plazas, A. A., & Bernstein, G. 2012, PASP, 124, 1113

Plazas, A. A., Bernstein, G. M., & Sheldon, E. S. 2014, PASP, 126, 750

Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, MNRAS, 425, 1951

Rhodes, J., Refregier, A., & Groth, E. J. 2000, ApJ, 536, 79

Roodman, A., Reil, K., & Davis, C. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9145, 914516

Rowe, B. 2010, MNRAS, 404, 350

Rowe, B., Jarvis, M., Mandelbaum, R., et al. 2015, Astronomy and Computing, 10, 121

Rozo, E., Rykoff, E. S., Abate, A., et al. 2015, ArXiv e-prints, arXiv:1507.05460

Rykoff, E., et al. 2015, in preparation

Sánchez, C., Carrasco Kind, M., Lin, H., et al. 2014, MNRAS, 445, 1482

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002, A&A, 389, 729

Schrabback, T., Erben, T., Simon, P., et al. 2007, A&A, 468, 823

Schrabback, T., Hartlap, J., Joachimi, B., et al. 2010, A&A, 516, A63

Seitz, C., & Schneider, P. 1997, A&A, 318, 687

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Sheldon, E. S. 2014, MNRAS, 444, L25

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Soumagnac, M. T., Abdalla, F. B., Lahav, O., et al. 2015, MNRAS, 450, 666

Tarlé, G., Bigelow, B., Boprie, D., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7739, 77393L

The Dark Energy Survey Collaboration. 2005, ArXiv Astrophysics e-prints, astro-ph/0510346

The Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2015, ArXiv e-prints, arXiv:1507.05552

Tyson, J. A., Roat, C., Bosch, J., & Wittman, D. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 394, Astronomical Data Analysis Software and Systems XVII, ed. R. W. Argyle, P. S. Bunclark, & J. R. Lewis, 107

Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

Velander, M., Kuijken, K., & Schrabback, T. 2011, MNRAS, 412, 2665

Voigt, L. M., & Bridle, S. L. 2010, MNRAS, 404, 458

von der Linden, A., Allen, M. T., Applegate, D. E., et al. 2014, MNRAS, 439, 2

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143

Zhang, J., & Komatsu, E. 2011, MNRAS, 414, 1047

Zuntz, J., Kacprzak, T., Voigt, L., et al. 2013, MNRAS, 434, 1604

[1]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
[3]Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[4]Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland
[5]Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA
[6]Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA
[7]Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA
[8]Institut de Física d'Altes Energies, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain
[9]Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
[10]Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
[11]Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany
[12]Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
[13]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
[14]Faculty of Physics, Ludwig-Maximilians University, Scheinerstr. 1, 81679 Munich, Germany
[15]Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany
[16]Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[17]Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
[18]Department of Physics, The Ohio State University, Columbus, OH 43210, USA
[19]SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[20]Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA
[21]Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile
[22]CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France
[23]Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France
[24]Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
[25]Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
[26]Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
[27]Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA
[28]National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA
[29]Institut de Ciències de l'Espai, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, 08193 Bellaterra, Barcelona, Spain
[30]George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
[31]Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[32]Australian Astronomical Observatory, North Ryde, NSW 2113, Australia
[33]Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, CEP 05314-970, São Paulo, SP, Brazil
[34]Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA
[35]Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain
[36]Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, UK
[37]Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
[38]Department of Physics, University of Illinois, 1110 W. Green St., Urbana, IL 61801, USA

This 2-column preprint was prepared with the AAS LaTeX macros v5.2.