The DES Science Verification weak lensing shear catalogues

Article (Published Version)

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **460,** 2245–2281 (2016)
Advance Access publication 2016 May 1

doi:10.1093/mnras/stw990

The DES Science Verification weak lensing shear catalogues

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,[10,11] D. Gruen,[14,15] 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,[16,22] S. Allam,[12] J. Annis,[12] A. Benoit-Lévy,[16] E. Bertin,[23,24] D. Brooks,[16] E. Buckley-Geer,[12] D. L. Burke,[7,19] D. Capozzi,[25] A. Carnero Rosell,[26,27] M. Carrasco Kind,[28,29] J. Carretero,[30,8] F. J. Castander,[30] J. Clampitt,[1] M. Crocce,[30] C. E. Cunha,[7] C. B. D'Andrea,[25] L. N. da Costa,[26,27] D. L. DePoy,[31] S. Desai,[10,11] H. T. Diehl,[12] P. Doel,[16] A. Fausti Neto,[26] B. Flaugher,[12] P. Fosalba,[30] J. Frieman,[12,32] E. Gaztanaga,[30] D. W. Gerdes,[9] R. A. Gruendl,[28,29] G. Gutierrez,[12] K. Honscheid,[17,18] D. J. James,[21] K. Kuehn,[33] N. Kuropatkin,[12] O. Lahav,[16] T. S. Li,[31] M. Lima,[34,26] M. March,[1] P. Martini,[17,35] R. Miquel,[36,8] J. J. Mohr,[10,11,14] E. Neilsen,[12] B. Nord,[12] R. Ogando,[26,27] K. Reil,[19] A. K. Romer,[37] A. Roodman,[7,19] M. Sako,[1] E. Sanchez,[38] V. Scarpine,[12] M. Schubnell,[9] I. Sevilla-Noarbe,[38,28] R. C. Smith,[21] M. Soares-Santos,[12] F. Sobreira,[12,26] M. E. C. Swanson,[29] G. Tarle,[9] J. Thaler,[39] D. Thomas,[25] A. R. Walker[21] and R. H. Wechsler[6,7,19]

Affiliations are listed at the end of the paper

Accepted 2016 April 25. Received 2016 April 18; in original form 2015 July 23

ABSTRACT

We present weak lensing shear catalogues 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 catalogues 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 catalogues for the full 5-yr DES, which is expected to cover 5000 square degrees.

Key words: gravitational lensing: weak – methods: data analysis – techniques: image processing – catalogues – surveys – cosmology: observations.

1 INTRODUCTION

Weak gravitational lensing provides a powerful statistical tool for studying the distribution of mass in the Universe. Light travelling from distant galaxies to Earth is deflected by the gravitational field

⋆ E-mail: michael@jarvis.net

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 per cent 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 per cent, 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. Nevertheless, 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 per cent. Future surveys may reach 0.1 per cent. This implies that systematic errors (i.e. biases) in the shape measurements need to be controlled at a level approximately three 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 atmosphere, 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 artefacts 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 subdominant 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, Squires & Broadhurst 1995; Rhodes, Refregier & Groth 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. Section 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; Bernstein 2010; Melchior et al. 2010; Voigt & Bridle 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 per cent for bulge+disc 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 per cent 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 S/N 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 modelling of additive systematic errors as proportional to PSF ellipticity, which we also apply as a calibration (cf. Section 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. Section 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. Van Waerbeke et al. 2000; Wittman 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 modelling 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. Section 7.1).

For removing problematic data, the CFHTLenS analysis trimmed the survey area to only those fields in which the shape catalogues 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. Section 5.1).

In this paper, we present the shear catalogue for the DES Science Verification (SV) data, described in Section 2. We derive

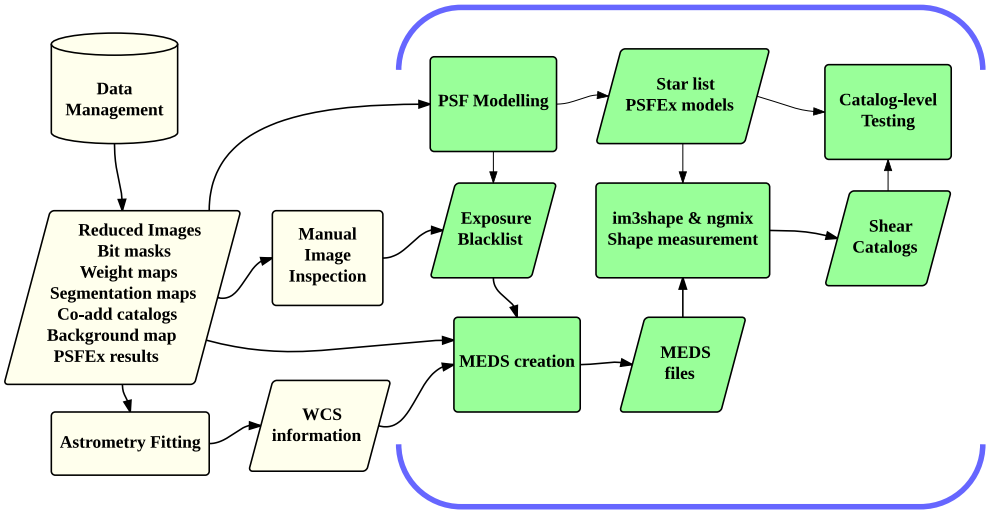

Figure 1. A flowchart showing the main stages in the production of the shear catalogues. The items inside the blue bracket are done by the weak lensing group in DES and are the principal subject of this paper.

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

2 DATA

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

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 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 2.7 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 DES officially started taking survey data in 2013 August (Diehl et al. 2014). It will cover about 5000 square degrees in the South Galactic Cap region, with ∼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 data set observed with the same instrument.

Before the start of the main survey, a small SV survey was conducted from 2012 November to 2013 February. 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-yr DES survey. The dither pattern matches that of the main survey, which uses large dithers to minimize the impact of any systematic errors related to the location on the field of view. Each tiling is typically observed on different nights to vary the observing conditions as much as possible. 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 were reduced by the DES Data Management (DESDM) system (Desai et al. 2012; Mohr et al. 2012), resulting in calibrated and background-subtracted images. Catalogues were produced using the software package Source Extractor (SEXTRACTOR; Bertin & Arnouts 1996; Bertin 2011). The PSF was characterized using the PSFEX package (Bertin 2011; for more details, see Section 4).

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

For weak lensing, we used 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 were important for determining a good set of galaxies to use for shear measurement.

[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 (Diehl et al. 2014), but it was still functional for the work presented here.

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

2.1 Object catalogue

The starting point for our object catalogue was the 'SVA1 Gold Catalogue',[4] which excludes regions of the data that were found to be problematic in some way, due to imaging artefacts, scattered light, failed observations, etc. The selection criteria for the Gold Catalogue included the following.

(i) Required object to have been observed at least once in each of the g, r, i, and z bands.
(ii) Required declination to be north of $61^{\circ}S$ to avoid the Large Magellanic Cloud and R Doradus, where the photometric calibration was found to have severe problems.
(iii) Removed regions with a high density of objects with 'crazy colours', 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.
(iv) Removed regions with a density less than 3σ below the mean density.
(v) Removed regions near bright stars. We eliminated a circular region around all stars detected in the Two-Micron All Sky Survey (2MASS; Skrutskie et al. 2006) brighter than $J_{\rm M}=12$ with a mask radius of $r=(-10J_{\rm M}+150)$ arcsec up to a maximum radius of 120 arcsec.
(vi) Removed 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 per cent of such objects fall into 4 per cent of the total area, so we removed all objects in that 4 per cent 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 brings this down to 148 square degrees for the Gold Catalogue.

The selection criteria listed above removed galaxies in a non-random way that varied across the sky. We characterized 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 Fig. 2. The white background represents the Gold Catalogue area. The coloured intensity represents the galaxy number density in the NGMIX catalogue (cf. Section 7.4).

The region used for the weak lensing analysis is somewhat smaller than the full Gold Catalogue region, because we additionally excluded CCD images with poor astrometric solutions (cf. Section 2.3), poor PSF solutions (cf. Section 4.2), and blacklisted CCDs containing bright stars, ghosts, airplanes etc. (cf. Section 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 catalogues of 139 square degrees. The intensity map for the IM3SHAPE catalogue (cf. Section 7.3) looks qualitatively similar, although it is about 40 per cent shallower (cf. Section 9.3).

[4] http://des.ncsa.illinois.edu/releases/sva1

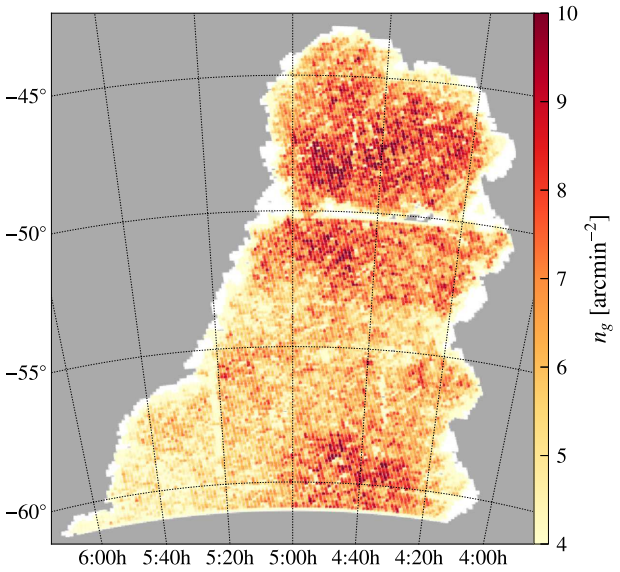

Figure 2. A HEALPIX map of the SPT-E region. The white background shows the full 'Gold' area. The colours show the galaxy density in the NGMIX shear catalogue (The map for IM3SHAPE is qualitatively similar, although about 40 per cent shallower.). The map has HEALPIX resolution `nside` = 512.

2.2 Galaxy selection

The preliminary galaxy selection was performed using standard SEXTRACTOR outputs from the i-band detections in the Gold Catalogue. The selection, in pseudo-code, was

```
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 were identified by the standard SEXTRACTOR classifier (`bright_test`). Fainter objects were considered stars if they were near the stellar locus in the `SPREAD_MODEL` measure, introduced by Desai et al. (2012), which uses a model of the local PSF to quantify the difference between PSF-like objects and resolved objects (`locus_test`; see also Bouy et al. 2013).

Objects whose best estimate of the total magnitude is much brighter than their PSF magnitude (`faint_psf_test`) are often spurious detections and were considered 'junk' in this classification. Our initial galaxy selection then included every object not classified as either a star or junk by this scheme.

Many faint stars and spurious detections remained in the catalogue at this stage. Later, further selection criteria, described in Section 9.1, were applied based on measurements from the shear pipelines. The initial selection was not intended to produce an accurate galaxy catalogue, but rather to produce a superset of the objects that would eventually be trimmed based on more stringent selection criteria. All objects in this preliminary galaxy catalogue were processed by both shear measurement algorithms (cf. Section 7).

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 milliarcsec 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 were assumed to take the form

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

where P is a gnomonic projection from the (curved) sky on to 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,[5] C is the 'instrument' solution, and E is the 'exposure' solution.

The instrument solution C was derived as follows. We took a series of ≈20 exposures of a rich star field in succession, with the telescope displaced by angles ranging from 10 arcsec up to the field of view of the camera. Coordinates of stars were determined in the pixel coordinates of each exposure, and we adjusted the parameters of the map in equation (1) to minimize the internal disagreement between sky coordinates of all the observations of each star. The solution also minimized the discrepancies between the positions of stars in the 2MASS (Skrutskie et al. 2006) catalogue 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 were left free for each of the 61 functional CCDs (cf. Diehl et al. 2014; Flaugher et al. 2015). While fitting the star field data, we forced all CCDs in a given exposure to share a common affine map E, so there were six additional free parameters in the fit for each exposure. The instrument maps C derived in this way were assumed to apply to all SV exposures taken with the same CCD in the same filter. The process was repeated for each of the g, r, i, z, Y filters.

Repeating the above 'star flat' procedure every few months revealed small changes in the astrometric map, consistent with rigid motion of some CCDs relative to the others at a level of ≈10 milliarcsec, probably occurring when the camera was cycled to room temperature for occasional engineering tasks.

To account for these small shifts, we determined an independent E function (6 degrees of freedom) for each CCD image in the SV data in another round of fitting. In this second stage, we minimized the disagreements between positions reported for all CCDs that contribute to each DES coadd image. The coefficients of the affine transformations E were allowed to float, but the higher order polynomials C were held fixed at the values determined from the star field data. These solutions again minimized residuals with respect to matching sources from the 2MASS catalogue 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 towards zenith, which are both properly included as part of the affine transformation E for each CCD. We did not, however, make any attempt to address the intraband 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 are consistent with errors in the astrometric maps of 10–20 milliarcsec rms in each coordinate. We found these errors to be coherent over arcminute scales in a given exposure, but were uncorrelated between distinct exposures. We interpret this to mean 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 the order of 10 milliarcseconds for our field of view and exposure time.

We found some remaining astrometric errors that were coherent over time and correlated with position on the detector array, which are consistent with small components of the electric fields transverse to the surface of the CCD in some places (Plazas, Bernstein & Sheldon 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 catalogues 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

$$\langle e \rangle = g \equiv \frac{\gamma}{1-\kappa}, \tag{2}$$

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 (2), where a and b are the semimajor and semiminor 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

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

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

$$I_{\mu\nu} = \frac{\int \mathrm{d}x\,\mathrm{d}y I(x,y)(\mu-\bar{\mu})(\nu-\bar{\nu})}{\int \mathrm{d}x\,\mathrm{d}y I(x,y)}. \tag{4}$$

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

While equation (2) is our goal for the shape estimates in our catalogue, it is inevitable that there will be systematic errors in the

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

shape measurements. A convenient parametrization, based on one first proposed by Heymans et al. (2006), uses a first-order expansion of the form,

$$\langle e \rangle = (1 + m) g_{\rm true} + \alpha e_{\rm PSF} + c, \quad (5)$$

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 four components if the leakage were anisotropic and involved cross terms,[6] 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 (5).

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):

$$\xi_+(\theta) = \left\langle e^*(\boldsymbol{x}) e(\boldsymbol{x} + \boldsymbol{\theta}) \right\rangle \quad (6)$$

$$\xi_-(\theta) = \langle e(\boldsymbol{x}) e(\boldsymbol{x} + \boldsymbol{\theta}) \exp(-4i \arg(\boldsymbol{\theta})) \rangle \,, \quad (7)$$

where * indicates complex conjugation.

Substituting equation (5) 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

$$\delta\xi_i(\theta) \simeq 2m\xi_i(\theta) + \alpha^2 \xi_i^{pp}(\theta) + \xi_i^{cc}(\theta) \quad (8)$$

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 autocorrelation functions of the PSF shapes, and ξ_i^{cc} are the autocorrelation 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 $8\,h^{-1}$ Mpc. Our requirement is that the systematic errors change the estimated value of σ_8 by less than 3 per cent, $\delta\sigma_8/\sigma_8 < 0.03$. 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

$$\delta\xi_i^{\rm max} = \frac{\partial \xi_i}{\partial \sigma_8} \delta\sigma_8. \quad (9)$$

[6] 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.

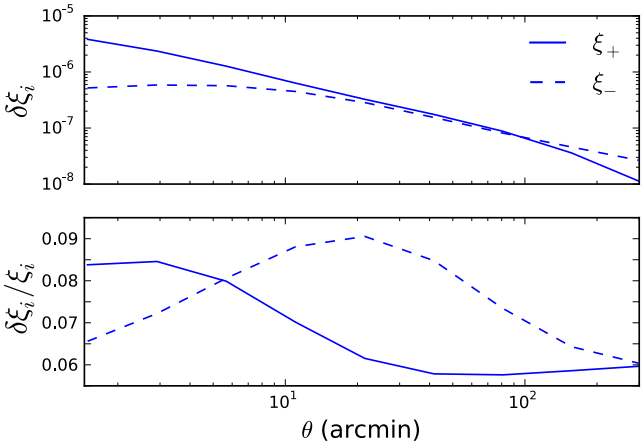

Figure 3. Requirement for the maximum systematic error contribution to the shear correlation functions. The blue lines correspond to $\delta\sigma_8/\sigma_8 = 0.03$ for each of the correlation functions ξ_+ (solid) and ξ_- (dashed). The top and bottom panels shows the requirement for the absolute and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

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

Fig. 3 shows the resulting requirements for $\delta\xi_i$ derived for a flat Λ cold dark matter (CDM) central cosmological model with $\sigma_8 = 0.82$, $\Omega_{\rm b} = 0.047$ and $\Omega_{\rm c} = 0.2344$, $h = 0.7$ and $n_{\rm s} = 0.96$.

We will apply this requirement to a number of different potential sources of systematic error. If each of them just barely pass the requirement, this would be a problem, since the total net systematic error would then exceed the requirement. We attempt to quantify the total realized systematic error in the shear measurements in Section 8.7; it is this total error that must be propagated into the next stage(s) of the analysis, along with any other non-measurement sources of systematic error (e.g. photometric redshift errors and intrinsic alignments) that may be relevant for each specific science application.

3.2 Multiplicative and additive errors

From equation (8), we find that the requirement on the multiplicative bias, m, is

$$|m| < \frac{1}{2} \left| \frac{\delta\xi_i^{\rm max}}{\xi_i} \right| . \quad (10)$$

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

$$|m| < 0.03. \quad (11)$$

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 $\sim$1 arcmin), the requirement comes from the zero-lag value of $\delta\xi_+$ in Fig. 3, $\langle c^2 \rangle < \delta\xi_+^{\rm max}(0)$, or

$$c_{\rm rms} < 2 \times 10^{-3}. \quad (12)$$

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

$$|\xi_i^{cc}(\theta)| < \delta\xi_i^{\rm max}(\theta) \quad (13)$$

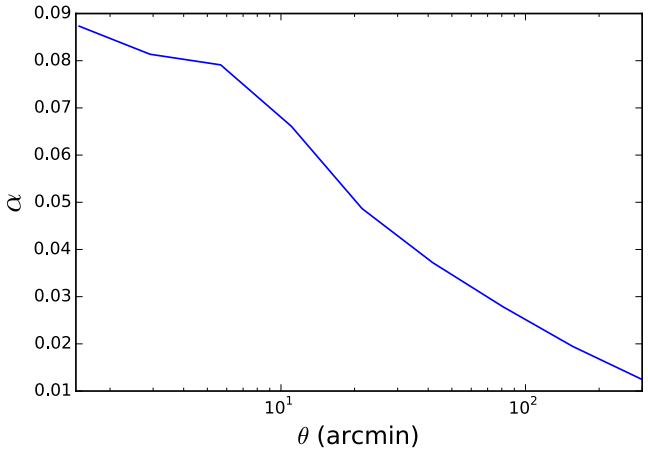

Figure 4. Requirement for the PSF leakage factor α based on the relative error in σ_8 being less than 3 per cent.

using the function shown in Fig. 3. The most notable example of this will be systematic effects due to the PSF: both leakage and modelling 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 (13) should ideally be satisfied for scales $\theta < 100$ arcmin, 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 (5) can be obtained from the general requirement on additive errors, equation (13),

$$\alpha^2 \xi_i^{pp}(\theta) < \delta\xi_i^{\rm max}(\theta), \tag{14}$$

which can be solved for α as

$$|\alpha| < \left(\frac{\delta\xi_i^{\rm max}(\theta)}{\xi_i^{pp}(\theta)}\right)^{1/2}. \tag{15}$$

Fig. 4 shows this requirement on α as a function of θ using the observed ξ_+^{pp} for DES SV data.[7] 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 Fig. 4 to determine a conservative requirement for α that would be applicable for scales $\theta <$ 100 arcmin:

$$|\alpha| < 0.03. \tag{16}$$

We will estimate α from the data in Section 8.2.

[7] See Fig. 20, top panels. We use the IM3SHAPE measurement of ξ_+^{pp} here.

3.4 PSF model errors

We now consider errors in the modelling 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):[8]

$$\delta e_{\rm sys} = (e - e_{\rm PSF})\left(\frac{T_{\rm PSF}}{T_{\rm gal}}\right)\frac{\delta T_{\rm PSF}}{T_{\rm PSF}} - \left(\frac{T_{\rm PSF}}{T_{\rm gal}}\right)\delta e_{\rm PSF}, \tag{17}$$

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

$$\begin{aligned}\delta\xi_+(\theta) = {} & 2\left\langle\frac{T_{\rm PSF}}{T_{\rm gal}}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right\rangle\xi_+(\theta) + \left\langle\frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle^2\rho_1(\theta) \\ & - \alpha\left\langle\frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle\rho_2(\theta) + \left\langle\frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle^2\rho_3(\theta) \\ & + \left\langle\frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle^2\rho_4(\theta) - \alpha\left\langle\frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle\rho_5(\theta),\end{aligned} \tag{18}$$

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

$$\rho_1(\theta) \equiv \left\langle\delta e_{\rm PSF}^*(\boldsymbol{x})\delta e_{\rm PSF}(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle \tag{19}$$

$$\rho_2(\theta) \equiv \left\langle e_{\rm PSF}^*(\boldsymbol{x})\delta e_{\rm PSF}(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle, \tag{20}$$

and we introduce three new statistics defined as[9]

$$\rho_3(\theta) \equiv \left\langle\left(e_{\rm PSF}^*\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x})\left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle \tag{21}$$

$$\rho_4(\theta) \equiv \left\langle\delta e_{\rm PSF}^*(\boldsymbol{x})\left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle \tag{22}$$

$$\rho_5(\theta) \equiv \left\langle e_{\rm PSF}^*(\boldsymbol{x})\left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle. \tag{23}$$

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

[8] 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 none the less a useful model for describing PSF modelling errors.

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

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

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 first term in equation (18) is a multiplicative systematic, so the relevant requirement comes from equation (11). We approximate the ensemble average as a product of two averages to set a requirement on the mean error in the PSF size

$$\left|\left\langle \frac{\delta T_{\rm PSF}}{T_{\rm PSF}} \right\rangle\right| < 0.03 \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^{-1}. \quad (24)$$

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 (13) are that each term be less than $\delta\xi_+^{\rm max}(\theta)$:

$$|\rho_{1,3,4}(\theta)| < \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^{-2} \delta\xi_+^{\rm max}(\theta) \quad (25)$$

$$|\rho_{2,5}(\theta)| < |\alpha|^{-1} \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^{-1} \delta\xi_+^{\rm max}(\theta). \quad (26)$$

We will test these requirements for our PSF model below in Section 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 catalogue keeps more small galaxies than the IM3SHAPE selection. We use the NGMIX value in Section 4.4, as it gives the more stringent requirement. For α, we conservatively use the value 0.03. We will find in Section 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 must be addressed in order to measure accurate shears is the convolution of the galaxy surface brightness profiles by the 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, but varies both spatially over the focal plane and temporally from one exposure to the next. The atmospheric component varies approximately according to a Kolmogorov turbulent spectrum. The optical aberrations have characteristic patterns due to features in the telescope optics.

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 selected the stars to be used to constrain the PSF model. As stars are point sources, observations of them provide a sample image of the PSF at the location of each star. We desired a high-purity sample of fairly bright stars to make sure we did 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, for some CCD images, the sets of objects identified as stars by the Modest Classification scheme[10] included a relatively high number of galaxies, and in other cases too few stars were identified. The cause of these failures is dependent on many factors, but may be partly related to the use of coadd data for the classification. The coadd PSF can change abruptly at the locations of chip edges in the original single-epoch images, which may have affected the stellar classification near these discontinuities.

Ultimately, the problems with the modest classifier were common enough that we decided to develop a new algorithm tailored specifically to the identification of a pure set of PSF stars. 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-fitting elliptical Gaussian profile using an adaptive moments algorithm. We found that these measures produce a flatter and tighter stellar locus than the `FLUX_RADIUS` value output by SEXTRACTOR, and is thus better suited for selection of stars. As a further improvement, we 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.

The stars are easily identified at bright magnitudes as a locus of points with constant size nearly 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 mag (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 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.

Fig. 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 for PSF modelling, even ignoring the inevitable few galaxies that get misidentified as stars. The CCDs on the DECam 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

[10] Stars were identified as `(bright_test OR locus_test)` in terms of the pseudo-code presented in Section 2.2

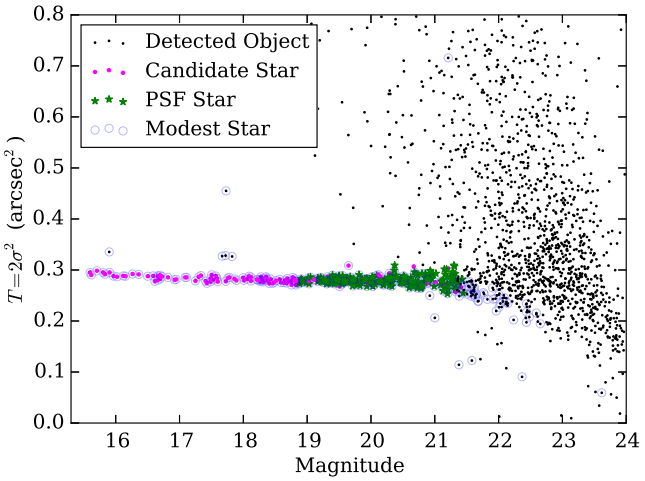

Figure 5. An example size–magnitude diagram for a single CCD image, used to identify stars. The size $T = 2\sigma^2$ is based on the scale size of the best-fitting elliptical Gaussian. The pink and green points are the objects initially identified as stars. The green points are the ones that pass our selection criteria outlined in Section 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 alternate star classification, called the Modest Classification within DES, which was found not to work as well for our specific purpose.

affects the electric field lines near each 2 mm × 2 mm 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 PSF full width half-max (FWHM) separation of the outline of a tape bump. The tape bumps are relatively small, so this procedure excludes less than 0.1 per cent of the total area of the CCD, but removes a noticeable bias in the PSF model near the bumps.

Another problem we addressed with regards to star selection is the so-called '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 appear 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 used 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.

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 catalogues, we instead partially avoided the problem by removing the brightest stars from our sample. Specifically, we removed all stars within 3 mag of the saturation limit for the exposure. That is, in our final selection of PSF stars we required that the brightest pixel in the stellar image be less than 6 per cent of the pixel full well. Since the brighter fatter effect scales approximately linearly with flux, this reduces the magnitude of the effect by a factor of 16. We were 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 the more sophisticated solution can be implemented.

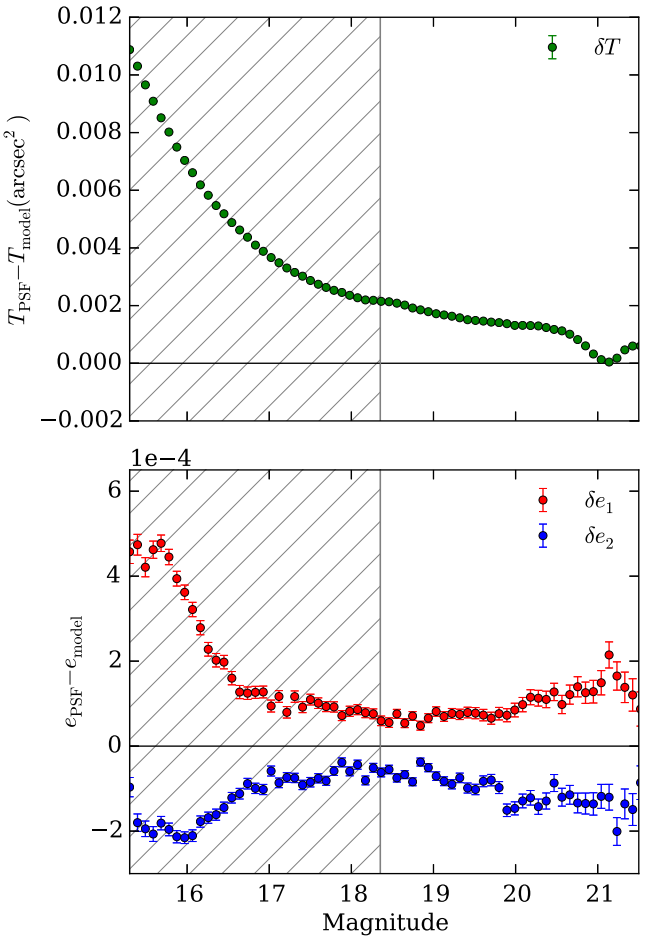

Figure 6. The residual size (top) and shape (bottom) of stars relative to that of the PSF model as a function of magnitude. The hatched region on the left shows the magnitude range of the stars we exclude from the sample to reduce the impact of the brighter fatter effect.

In Fig. 6, we show 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 Section 4.3. For the measurements of the sizes and shapes described here, we used the implementation of the HSM (Hirata & Seljak 2003; Mandelbaum et al. 2005) algorithm included in the GALSIM software package. The hatched region marks the range we excluded to avoid the spurious increase in PSF size from the brighter fatter effect. In Fig. 6, we have also shown the mean difference in ellipticity due to the brighter fatter effect; it affects the shapes of the stars as well as the size.

We do not yet understand why the residual sizes and shapes shown in Fig. 6 do not level off to zero at fainter magnitudes where the brighter fatter effect is negligible. The requirement on this residual value is given by equation (24). 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.

In the complete process described above, we find a median of 130 useful stars per CCD image, which we use to constrain the PSF model. The distribution is shown in Fig. 7.

In Fig. 8, we show the distribution of the median measured FWHM for the PSF stars used in our study, restricted to the exposures used for shear measurements. The overall median seeing

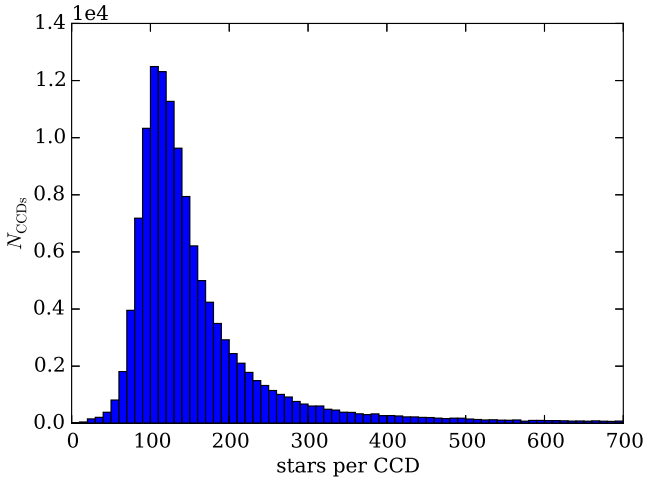

Figure 7. The distribution of the number of stars per CCD image used for constraining the PSF model.

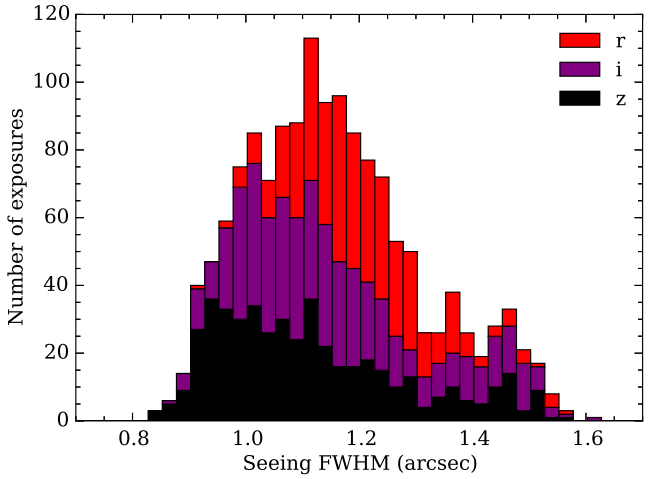

Figure 8. The distribution of the median FWHM of the PSF stars in the non-blacklisted exposures (cf. Section 5.1). The *r*-, *i*- and *z*-band exposures are indicated from top to bottom within each bar in red, purple, and black.

of these exposures was 1.08 arcsec. The *r*, *i* and *z* bands had median seeing of 1.11, 1.08, and 1.03 arcsec, respectively. This was somewhat worse than expected (0.9 arcsec) and reflects the fact that a number of problems related to the instrument, telescope, and control software were being diagnosed and fixed concurrently with the observations. The realized seeing has significantly improved in the subsequent main survey observations (Diehl et al. 2014).

For some CCD images no stars passed our final selection criteria, usually 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 zero or very few objects detected. In less extreme cases, a bright object can sufficiently contaminate the fluxes and sizes of the other detections that the stellar locus is either difficult to find or merges with the galaxy locus at a fairly bright magnitude, such that the brighter fatter cut excludes the entire sample.

Whenever the process failed for any reason on a given CCD image, we flagged the image and excluded it from being used in subsequent shear estimation. We also flagged 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 were added to the set of blacklisted images described in Section 5.1.

4.3 PSF measurement and interpolation

To measure the PSF and its spatial variation on each CCD, we used 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 Section 4.2, we removed some of the stars in the catalogue to avoid the brighter fatter effect and the tape bumps. This edited catalogue of stars was then passed to PSFEX.

We used 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 forced 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 arcsec on a side.

For the interpolation, we used a 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 examined the differences between the measured shapes (using the HSM algorithm again) of actual stars on the image and the corresponding values for the PSFEX model at the locations of these stars. In Fig. 9, we show the 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 Section 4.4, and we will show 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 modelling 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 interpolate the PSF in sky coordinates rather than pixel coordinates, so that WCS variations can be modelled separately from 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 Section 3.4. If the PSF errors were random, independent values for each galaxy, they would constitute merely an additional

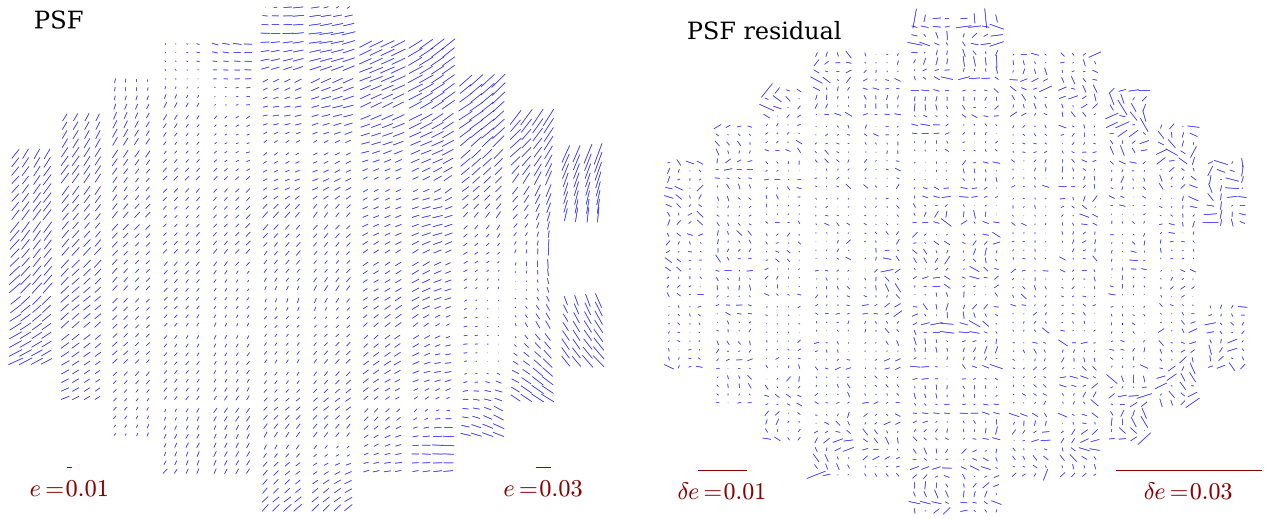

Figure 9. Whisker plots of the mean PSF pattern (left) and of the mean residual after subtracting off the model PSF (right) as a function of position in the focal plane. The length of each whisker is proportional to the measured ellipticity, and the orientation is aligned with the direction of the ellipticity. 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 per cent 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.

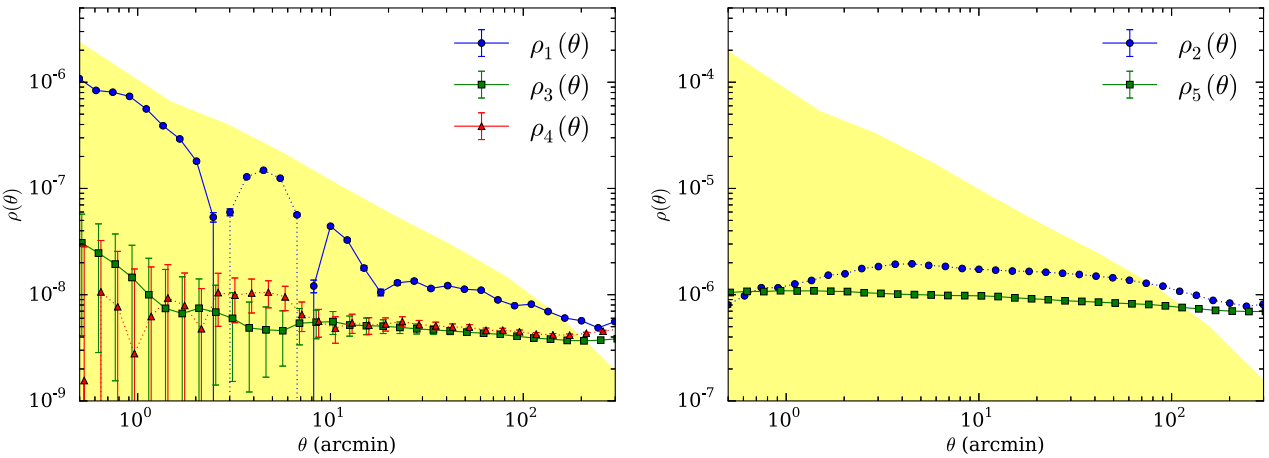

Figure 10. The ρ statistics for the PSF shape residuals. Negative values are shown in absolute value as dotted lines. The shaded regions are the requirements for SV data.

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 (18).

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 Section 3.4,

$$\rho_1(\theta) \equiv \left\langle \delta e^*_{\rm PSF}(\boldsymbol{x}) \delta e_{\rm PSF}(\boldsymbol{x}+\boldsymbol{\theta}) \right\rangle \tag{27}$$

$$\rho_2(\theta) \equiv \left\langle e^*_{\rm PSF}(\boldsymbol{x}) \delta e_{\rm PSF}(\boldsymbol{x}+\boldsymbol{\theta}) \right\rangle , \tag{28}$$

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 (21)–(23) and are generally smaller than the two described by Rowe (2010).

Fig. 10 shows the results for these five statistics. The shaded regions show the requirements, from equations (25) and (26). In all cases, the results are seen to be passing our requirements for scales less than about 100 arcmin. We see in Fig. 10 that ρ_1 changes sign twice and is below the requirement line by only a factor of $\sim$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 1 per cent.

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 Section 2, we used the coadd images for object detection and deblending. For shear measurement, we worked directly with the pixel data from the original single-epoch images (cf. Section 7.1). To simplify the bookkeeping we developed a new data storage format, which we named MEDS.[11]

We created a MEDS file corresponding to each coadd image. In these files, we stored 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 were 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 were sky-subtracted and then scaled to be on a common photometric system, which simplified the model fitting using these images. We also stored the local affine approximation of the WCS function, evaluated at the object centre, so that models could be made in sky coordinates and constrained using the different image coordinates for each postage stamp.

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

5.1 Exposure selection

We did not use all single-epoch images for measuring shapes. We excluded a small fraction of the CCD images that had known problems in the original data or in some step of the data reduction and processing. We created simple 'blacklist' files, in which we stored information for CCD images we wished to exclude, and that information was incorporated into the MEDS files as a set of bitmask flags. Postage stamps from blacklisted images were then easily excluded from the analysis when measuring shears. Here, we list some of the reasons that images were blacklisted.

Some of the astrometry solutions (cf. Section 2.3) provided a poor map from CCD coordinates to sky coordinates. This happened primarily 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. Section 4) provided a poor model of the PSF across the CCD. In some cases, there were too few stars detected to constrain the model; occasionally there was some error when running either the star finding code or PSFEX.

A small fraction of the SV images were contaminated by bright scattered-light artefacts. Scattered-light artefacts 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 Catalogue (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 2013 April, filter baffles were installed to block some of this scattered light, and non-reflective paint was applied to the filter changer and shutter in 2014 March (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 DECam 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. We identified these airplane trails by eye and removed a total of 56 individual CCD images due to airplane contamination (corresponding to four 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 s 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 colour of objects that they cross. For SV, the 'crazy colours' cut mentioned in Section 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. For this analysis, we used what we call an 'überseg' mask, constructed from the weight maps, segmentation maps and locations of nearby objects.

To create the überseg mask, we started with the SExtractor segmentation map from the coadd image, mapping it on to the corresponding pixels of the single-epoch images. We prefer this map to 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 were either associated with other objects in the segmentation map or were closer to any other object than to the object of interest. The result was 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 Fig. 11. In tests on a simulation with realistically blended galaxies (cf. Section 6.2), we found a large reduction in the shear biases when using the überseg masking as compared to the ordinary SExtractor segmentation maps. In particular, when using ordinary segmentation maps we found a significant bias of the galaxy shape in the direction towards neighbours. With the überseg masking, such a bias was undetectable.

6 SIMULATIONS

Simulations were a crucial part of our shear pipeline development process, providing data with known values of the applied shear for

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

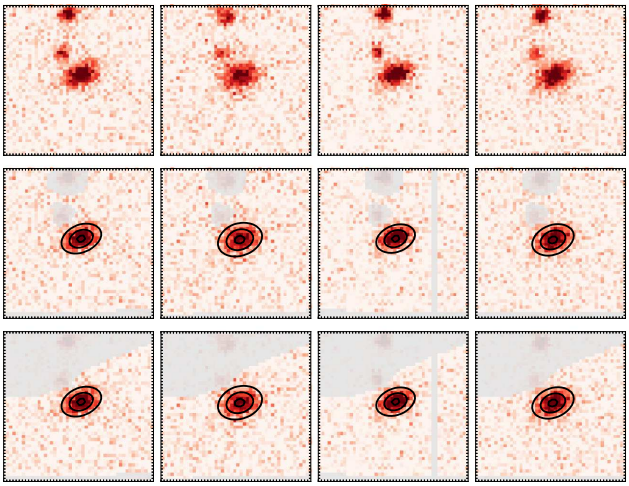

Figure 11. Example galaxy image demonstrating two masking strategies. The top row shows the original postage stamps in the MEDS file. The second row shows the result when only the SExtractor segmentation map was used to mask neighbours. The third row shows the result when the überseg algorithm was used to mask neighbours, as described in the text.

testing the shear estimation code. There is no such absolute calibration source in the real data. In addition to many small targeted simulations designed to answer particular questions about the algorithms, we developed two general purpose simulations that we used extensively to test the shear pipelines.

The first, which we call GREAT-DES, was modelled on the GREAT3 challenge. We used the GREAT-DES simulation to test 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 were placed on postage stamps, so there were no blending or object detection issues to consider.

The second, which we call the end-to-end simulation, was a high S/N simulation with analytic galaxy models with elliptical isophotes. The motivation with these simulations was to test that various bookkeeping details were 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. In these simulations, we also tested the efficacy of the überseg masking (cf. 5.2).

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

We used the GREAT-DES simulation to test the precision and accuracy of our shear measurement codes, using DES-tuned 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 ignored issues of crowding, bad pixels, and imaging artefacts, but we otherwise attempted to make the images a close approximation to the DES SV data.

We built the GREAT-DES simulation using galaxies from the COSMOS survey (Koekemoer et al. 2007), made available for use with GALSIM.[12] Kannawadi, Mandelbaum & Lackner (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 started with the entire COSMOS sample distributed for use with GALSIM and discarded objects that were flagged as unusable in the GREAT3 challenge (Mandelbaum et al. 2014), which removed about 3 per cent of the objects and left more than 54 000 COSMOS galaxies available for use in the simulation. Next, we selected individual galaxies from this set in such a way as to mimic the distribution of galaxy properties found in DES SV data.

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

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

Each of the COSMOS galaxies was used hundreds of times, with different noise realizations, different random orientations and different centroid offsets. We did not use 90° 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. We instead randomly oriented each galaxy. With this choice, we retained the ability to select subsets of the galaxies according to their measured characteristics. Such selections tends to break up the pairs, which obviates any advantage from using them. More importantly, additive errors can cancel between the pairs of galaxies, which would hide important systematic errors that would yet appear real data. 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 could 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, we desired the simulation to be closely representative of the data (see e.g. Bergé et al. 2013; Bruderer et al. 2016). To check that we have achieved this goal, we compared the IM3SHAPE measurements of relevant galaxy properties in GREAT-DES to those in the SV data. In Fig. 12, we show 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, shows

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

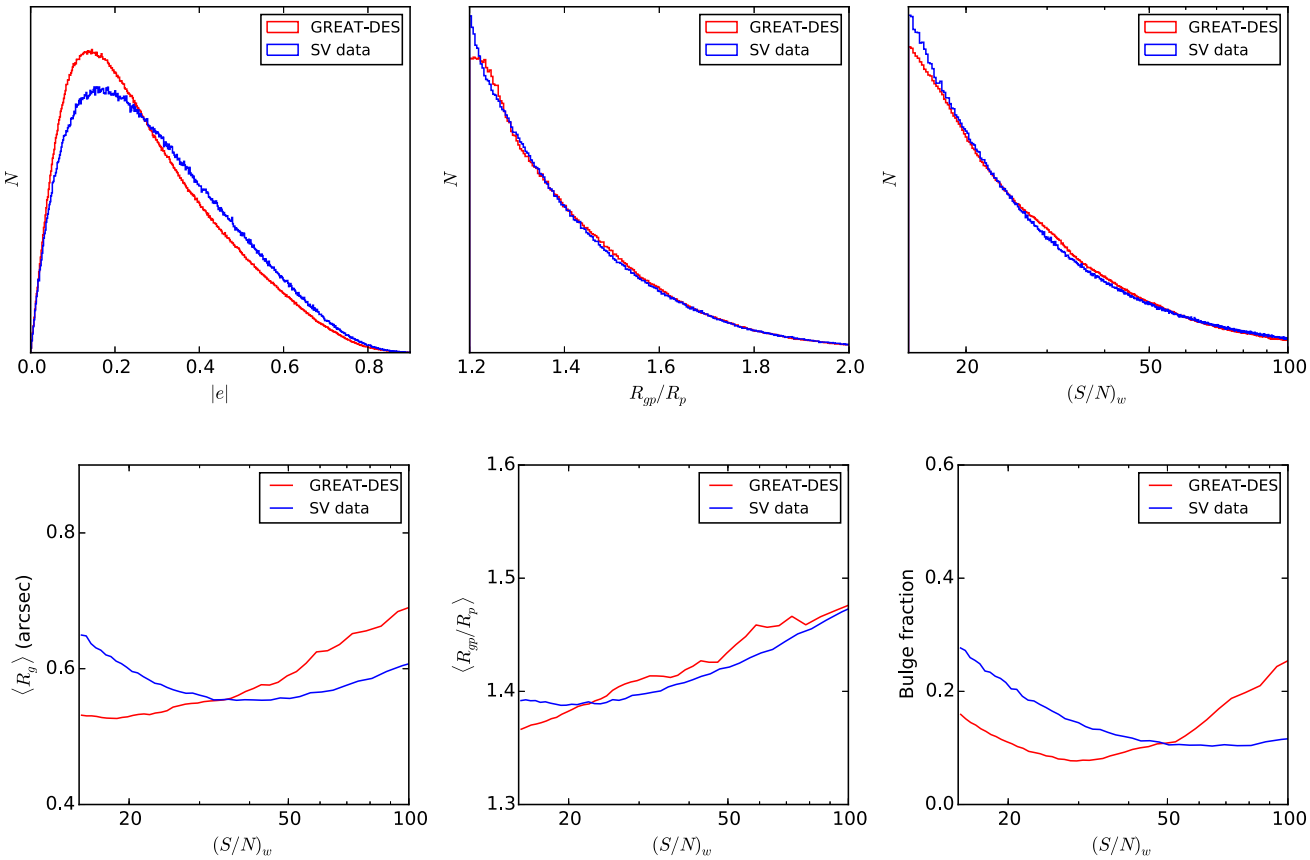

Figure 12. A comparison of the galaxy properties in GREAT-DES (red) and the SV data (blue). The top row shows histograms of $|e|$, R_{gp}/R_p, and S/N as measured by IM3SHAPE. The bottom row shows the dependence of $\langle R_g \rangle$, $\langle R_{gp}/R_p \rangle$, and bulge fraction as functions of S/N.

systematic disagreement as a function of S/N, 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 Fig. 12 is arbitrary; the analogous plot of NGMIX measurements shows similar agreement, except that there is no estimate of bulge fraction from the NGMIX exponential disc 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 concentration of the galaxies.

Since NGMIX uses an exponential disc model (cf. Section 7.4.1), and thus has worse model bias for bulge-like galaxies, the discrepancy in the bulge fraction limits our ability to infer what the model bias would be in real data (cf. Section 8.5). The IM3SHAPE shear estimate is based on a bulge-or-disc galaxy model (cf. Section 7.3.1), which has less overall model bias (Kacprzak et al. 2014), so the discrepancy may have less impact. However, the ability to accurately choose bulge versus disc is dependent on S/N, so our ability to test this aspect of the fitting is also somewhat limited in GREAT-DES.

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 used simple exponential disc profiles, which have elliptical isophotes when sheared, and the images were rendered with relatively little pixel noise. The fundamental shape estimation problem is thus straightforward for both algorithms.

The starting point for this simulation was one of the actual MEDS files from the data, along with the corresponding coadd catalogue, 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 built new versions of these single-epoch images using exponential disc galaxies with the same size, flux, ellipticity, and celestial position as the measurements of the real galaxies. We used variable elliptical Gaussian profiles for the PSF, using different parameters for each single-epoch image. The convolved images were rendered at the correct position on each image using the original WCS. With GALSIM, we applied the Jacobian of the WCS to the surface brightness profile as well, so this important detail was handled correctly. Objects that were deemed to be stars in the original catalogue (based on the SEXTRACTOR SPREAD_MODEL being less than 0.003) were drawn as a PSF profile, with the same flux as the original object. Finally, we added the original background sky level to the image, but with relatively small noise so that the faintest galaxies had $S/N > 200$.

We then ran these images through the full weak lensing pipelines, 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 were then compared to the true shapes of the simulated galaxies, which were 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. Section 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 neighbour on the same postage stamp, light from the neighbour 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 neighbour.

Switching to the überseg mask made a big difference; we found the measured shapes were then much closer to the true values. We found there was still a small effect from neighbours, which amounted to a slight increase in the effective shape noise for such objects, but we no longer detected any systematic bias in the shape estimates due to unmasked flux from neighbouring objects.

7 SHEAR MEASUREMENT

We used two different shear measurement codes for this study: im3shape and ngmix, both of which are based on model-fitting. im3shape performs a maximum likelihood fit using a bulge-or-disc galaxy model (cf. Section 7.3.1). ngmix uses an exponential disc model, exploring the full N-dimensional posterior likelihood surface with an informative prior applied on the ellipticity (cf. Section 7.4.1).

With both shear methods we used the psfex models of the PSF detailed in Section 4.3, although the way the PSF model was used differed. psfex produces a 2D image of the PSF profile at the location of each galaxy. With im3shape, we resampled the PSF image to a higher resolution grid and performed the convolution with the galaxy model via fast Fourier transform (FFT). With ngmix, we fit three free elliptical Gaussians to the PSF image and performed an analytic convolution with the galaxy model, which was also approximated as a sum of Gaussians, resulting in very fast model creation.

Finally, with both shear codes we used the MEDS files described in Section 5 to constrain the galaxy models, using pixel data from the original single-epoch images rather than using the coadd image, which we only used for object detection.

We discuss the details of the multi-epoch fitting process in the next section, Section 7.1. In Section 7.2, we define what we mean by *S/N*. The details of the im3shape and ngmix algorithms are given in Sections 7.3 and 7.4. Finally, our strategy for blinding the shear estimates is described in Section 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 complexity in 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 used the MEDS described in Section 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 was stored in the MEDS file and used during modelling.

For both codes (ngmix and im3shape), the model for a given set of galaxy parameters was generated in celestial coordinates. For ngmix, we modelled the PSF in celestial coordinates as well, and convolved it with the galaxy model analytically. We then compared this model to the observed data using the WCS transformation. For im3shape, we modelled the galaxy and PSF in image coordinates and convolved via FFT.

7.2 *S/N* ratio

Before we describe the algorithms we used 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 Section 7.3.2, where we discuss how im3shape calibrates the shear bias) and in later sections such as Section 8.3, where we test that the shear is independent of S/N, and Section 9.1, where we use S/N while selecting galaxies for the final shear catalogues.

There is no single definition for the S/N of an image or a surface brightness profile. Rather, an 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

$$S/N \equiv \frac{x}{\sqrt{\mathrm{Var}(x)}}. \tag{29}$$

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,

$$\hat{I}_w = \sum_p w_p I_p, \tag{30}$$

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.[13]

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-fitting model of the galaxy, which we call m_p, as part of the weight. The S/N of this statistic is thus estimated as

$$(S/N)_w = \frac{\sum_p m_p I_p / \sigma_p^2}{\left(\sum_p m_p^2 / \sigma_p^2\right)^{1/2}}. \tag{31}$$

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.

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

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 catalogue 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 an 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 (31), as the actual data are no longer appropriate for this counterfactual surface brightness profile. The 'roundified' S/N estimator is then

$$(S/N)_r = \frac{\sum_p m_p^r m_p^r/\sigma_p^2}{\left(\sum_p (m_p^r)^2/\sigma_p^2\right)^{1/2}} = \left(\sum_p (m_p^r)^2/\sigma_p^2\right)^{1/2}. \quad (32)$$

We find both measures of the *S/N* 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. Section 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 *S/N* of the actual galaxy than to that of a counterfactual round version of the galaxy. Thus, the 'noise bias' calibration in fact also approximately 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 catalogue.

7.3 Shear measurements with IM3SHAPE

IM3SHAPE is a maximum-likelihood model-fitting code, which we used to fit de Vaucouleurs bulge and exponential disc components to galaxy images. The code 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. Section 8). Previously each galaxy was modelled as the sum of two components, a bulge and a disc. In this paper, we fit each galaxy twice: once as a pure bulge and once as a pure disc. For shear estimation, we used the model with the higher likelihood, unless that model was flagged as a bad fit (cf. Section 7.3.3). If both models were flagged the galaxy was excluded from the catalogue.

We found this 'bulge or disc' scheme to be much more robust on simulations of realistic galaxies. This scheme produced good model fits in almost all cases, whereas with the previous 'bulge plus disc' scheme we frequently found the best-fitting model was unphysical, with highly negative-flux components.

The parameters of the best-fitting model were found using the numerical optimizer LEVMAR[14] (Lourakis 2004), which is an implementation of the Levenberg–Marquardt (LM) algorithm (Levenberg 1944; Marquardt 1963), iterating towards a model image which minimizes the χ^2 with the data image.

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

We made a number of additions to the original code presented in Zuntz et al. (2013); 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.[15] One particularly useful infrastructure improvement was the implementation of a PYTHON interface to the existing C functions. We used 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 Section 7.3.3 below.

The biggest change we made to IM3SHAPE was 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 centred 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 Section 7.1.

7.3.1 Bulge or disc model

Each galaxy model was defined by six varied parameters: the amplitudes of either the bulge or disc components (A_b, A_d), a centroid relative to the nominal detection position (δu, δv), 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 used the local affine approximation to the WCS for each postage stamp (stored in the MEDS file) to transform these parameters into each image's local pixel coordinate system. Schematically,

$$\left\{\delta u, \delta v, e_1^c, e_2^c, r^c\right\} \rightarrow \left[\left\{\delta x, \delta y, e_1^p, e_2^p, r^p\right\}_{\text{Image1}}, \left\{\delta x, \delta y, e_1^p, e_2^p, r^p\right\}_{\text{Image2}}, \ldots\right], \quad (33)$$

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.

Given the appropriate parameters for each postage stamp, we then built the galaxy models in pixel coordinates, each convolved by the correct PSF for that stamp, and computed 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 gave us the final likelihood to use for that set of model parameters. We then iterated to find the maximum likelihood parameters for each galaxy. The maximum likelihood was typically found in less than 50 iterations. At 150 iterations, we stopped the algorithm and declared failure.

[14] http://users.ics.forth.gr/~lourakis/levmar/
[15] https://bitbucket.org/joezuntz/im3shape

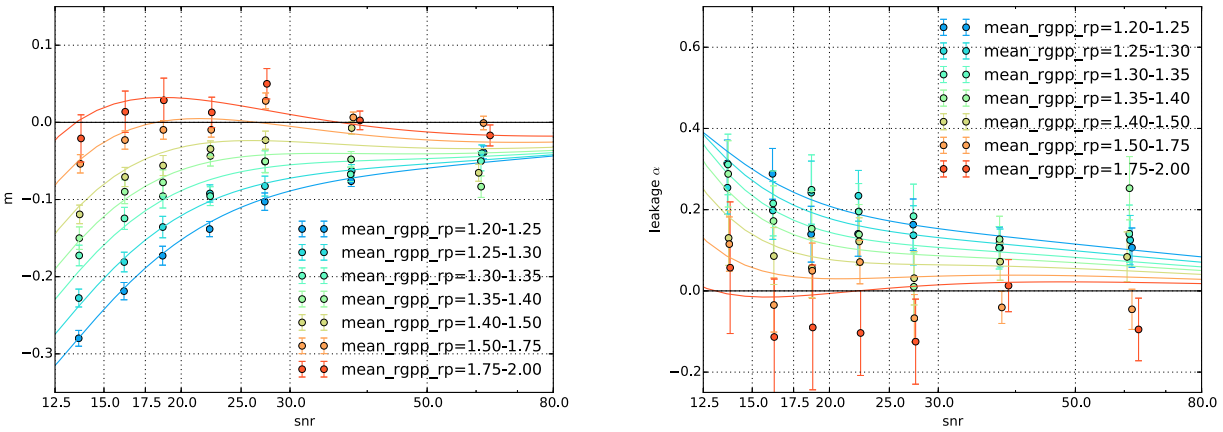

Figure 13. Shear bias for IM3SHAPE measurements on the GREAT-DES simulation: multiplicative bias (left) and PSF leakage (right), as functions of the measured $(S/N)_w$ and R_{gp}/R_p. The fits, which are used to calibrate the shear estimates on the data, are smooth functions in both of these variables. Solid lines show the fits versus $(S/N)_w$ at particular choices of R_{gp}/R_p.

The LM code that we used 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:

$$I_p^{\rm LM} = I_p/\sigma_p$$
$$m_p^{\rm LM} = m_p/\sigma_p. \quad (34)$$

This maintains the χ^2 per pixel that the LEVMAR algorithm uses as its objective function. The estimates of σ_p came 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 (Kacprzak et al. 2012; Refregier et al. 2012). The fitted model parameters are a non-linear 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 found a small *selection bias*, which is introduced by using recommended IM3SHAPE flags (cf. Section 7.3.3) and the selection based on galaxy size and S/N (cf. Section 9.1). We also expect a small amount of *model bias* due to realistic galaxies not always being well fit by our bulge-or-disc 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 calculated bias corrections of the form shown in equation (5). Specifically, we fit for m and α as functions of $(S/N)_w$ (defined in equation 31) 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. Section 6.1). We ran 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 used a regularization term in the least-square fit and optimized it such that the fitted surface has a reduced $\chi^2 = 1$. A similar procedure was applied to α, where we used 18 parameters in the fit. In Fig. 13, we show 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 checked 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 did not vary by more than 1 per cent.

In Section 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 used these fits to estimate the multiplicative and additive corrections to use for every galaxy in the IM3SHAPE catalogue. 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. Section 9.2).

Note that a 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 Section 8.5, we test that the shear calibration nevertheless 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 generated 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 showed evidence of imaging artefacts or some other problem that violates the assumptions we have made in the model, so we wanted to be able to remove these objects from the final shear catalogue.

We distinguished 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 B1. In Section 9.1, we will detail the final selection criteria that we recommend for the IM3SHAPE catalogue, 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 assigned 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:

$$w = \frac{1}{\sigma_{\rm SN}^2 + \sigma_e^2}. \quad (35)$$

The LEVMAR LM implementation we used produces an estimate of the parameter covariance for each galaxy as a by-product of optimization, but we did not use this estimate to give us weights, for two reasons. First, we found it to show rather wide scatter when compared to Markov Chain Monte Carlo (MCMC) tests, often showing spurious parameter covariances. Secondly, our physical parameters are a non-linear function of the numerical parameters in some regimes, as discussed in Zuntz et al. (2013).

To estimate the appropriate weight for each galaxy, we instead used the measured shears from the GREAT-DES simulation. We grouped galaxies in bins of $(S/N)_w$ and R_{gp}/R_p. We then measured 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 was taken, and the weight was then given by the inverse variance.

We also imposed a maximum weight set by the mean variance of all high-S/N bins. Otherwise spuriously low variance estimates in some sparsely populated bins resulted 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,[16] and is available for download.[17]

In NGMIX, both the PSF profile and the galaxy are modelled 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 disc model

For the galaxy model, NGMIX supports various options including exponential discs, 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 catalogues, we used the exponential disc model for galaxies.

Using this simple disc model resulted in detectable model bias (cf. Section 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 modelling in detail for future DES analyses.

We constructed the model in celestial coordinates and fit it to multiple epochs and bands simultaneously (cf. Section 7.1). The centre, size and ellipticity were set to be the same for all bands and epochs, but the flux was allowed to vary between bands. For this study, we combined bands r, i, z, resulting in eight free parameters:

(i) u_c, v_c, the object centre in celestial coordinates, relative to the fiducial centre from the coadd object catalogue. The units are arcseconds.
(ii) e_1, e_2, the ellipticity.
(iii) 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.
(iv) F_k, the flux in each of the r, i, z bands.

7.4.2 Image fitting

The NGMIX code supports multiple paradigms, all of which were used in the current analysis.

(i) Exploration of the full likelihood surface for a given set of model parameters with an 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 PSF.

(ii) Maximum-likelihood fitting using any of a variety of function minimizers. We used 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.

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

(iii) Expectation Maximization (EM; Dempster, Laird & Rubin 1977), fitting directly to the pixels only. This method is used for PSF fitting.

For PSF measurement, the EM code was used, with three completely free Gaussians. 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-centre PSF components that are occasionally found in the SV PSF images were fitted without instability.

We chose 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.[18]

(i) 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).

(ii) 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.

(iii) 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.

(iv) 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 disc model for galaxies (cf. Section 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,

$$\langle e_\mu \rangle = \frac{\int \mathcal{L}(\boldsymbol{e}) p(e) e_\mu \mathrm{d}\boldsymbol{e}}{\int \mathcal{L}(\boldsymbol{e}) p(e) \mathrm{d}\boldsymbol{e}} \tag{36}$$

$$\simeq \frac{\sum_j p(e^j) e^j_\mu}{\sum_j p(e^j)}, \tag{37}$$

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.

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

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

$$\simeq 1 - \left[\frac{\sum_j \left(\langle e_\mu \rangle - e^j_\mu \right) \frac{\partial p}{\partial e_\mu}}{\sum_j p(e^j)} \right]. \tag{38}$$

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:

$$g_\mu = \frac{\sum_i \langle e^i_\mu \rangle}{\sum_i s^i_\mu}, \tag{39}$$

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

$$w = \frac{1}{2\sigma^2_{\rm SN} + C_{1,1} + C_{2,2}}, \tag{40}$$

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. Section 7.4.4), and $C_{i,j}$ are elements of the 2×2 ellipticity subset of the covariance matrix produced by NGMIX.

The sensitivities in equation (38) 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 chose 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 (38).

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 Section 6.1 using a prior with intrinsic variance in ellipticity 35 per cent higher than the true variance, and found the multiplicative bias increased by (1.3 ± 0.2) per cent.

For application to real data, we based 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

$$p(e) = 2\pi e A \frac{\left(1 - \exp\left[\frac{e-1}{a}\right]\right)}{(1+e)\sqrt{e^2 + e_0^2}}\, c(e) \tag{41}$$

$$c(e) = \frac{1}{2}\left(1 + \mathrm{erf}\left[\frac{e_{\rm cut} - e}{\sigma_e}\right]\right). \tag{42}$$

This model is a modified version of that introduced in Miller et al. (2013). Note in particular the cutoff at high ellipticities achieved by

[18] 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.

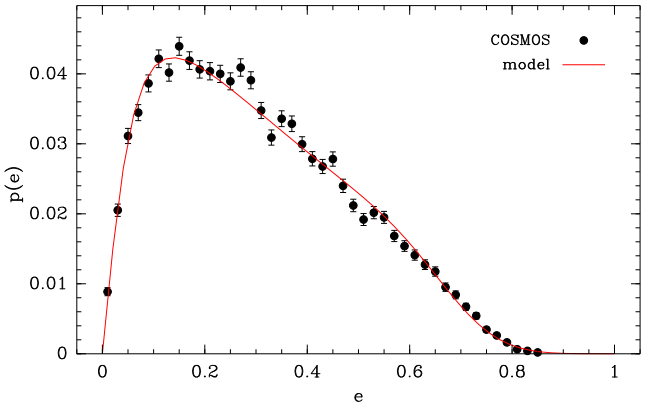

Figure 14. Distribution of shapes for COSMOS galaxies, selected as described in the text. The model fit was used as a prior for the NGMIX shear analysis.

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

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

using an error function. We found this formula improved the fit to the distribution of ellipticities.

We fit this model 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 Fig. 14. The best-fitting parameters are given in Table 1.

We used this same prior for all galaxies, but the distribution of COSMOS galaxy shapes depends on redshift. However, for this study we found that the uncertainties due to the redshift dependence of the shape distribution were subdominant to model bias for NGMIX (cf. Section 8.5).

7.5 Blinding

We blinded the shape catalogues from both pipelines before they were used for any tests or SV science papers. We did this to prevent the *experimenter bias* effect, wherein researchers work harder on finding bugs, 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).

We blinded the SV shear catalogues 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. We only unblined the catalogues after the analysis for a given paper was finalized.

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 developed an extensive test suite to check that the shear catalogues do not have significant systematic errors that would adversely affect weak lensing science. While there is no way to definitely prove that the shear catalogues 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 were formulated as 'null tests', which should have zero signal in the absence of systematic errors. Most of our null tests were similar to ones that have been performed in previous analyses (cf. e.g. Jarvis et al. 2003; Schrabback et al. 2010; Velander, Kuijken & Schrabback 2011; Heymans et al. 2012b; Jee et al. 2013; Kuijken et al. 2015).

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

(i) *Spatial tests* to check for systematic errors that are connected to the physical structure of the camera. Examples of these are 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 (Section 8.1).

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

(iii) *Galaxy property tests* to 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 (Section 8.3).

(iv) *B-mode statistics* to 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 direct test of systematic errors (Section 8.4).

(v) *Calibration tests* to 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 (Section 8.5).

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

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

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

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

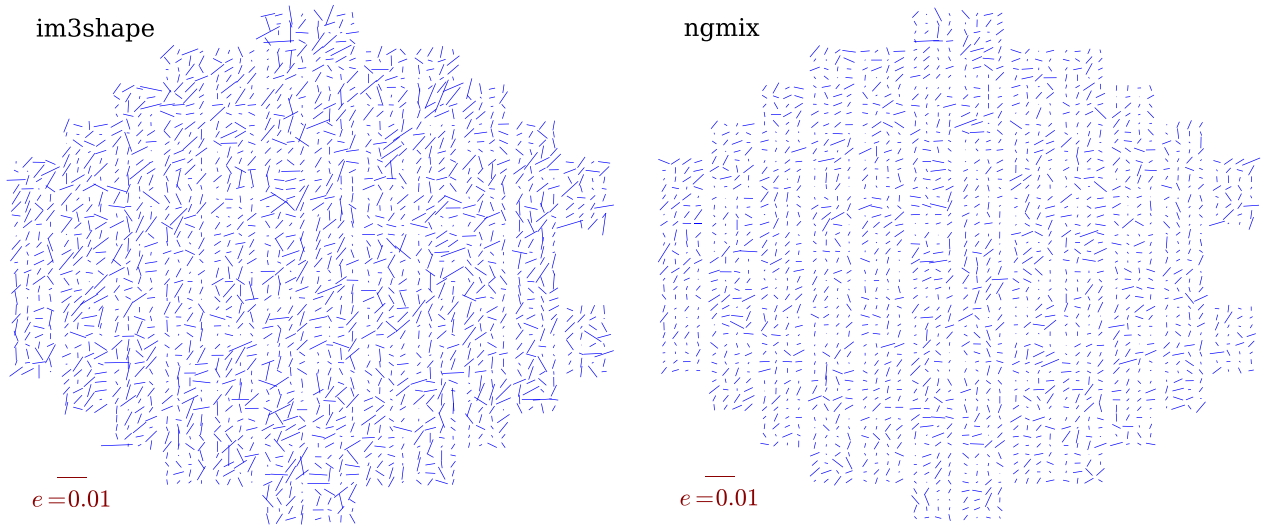

Figure 15. Whisker plots of the mean shear binned by position in the focal plane for IM3SHAPE (left) and NGMIX (right). As in Fig. 9, the length of each whisker is proportional to the mean shear, and the orientation is aligned with the direction of the ellipticity. A 1 per cent whisker is shown for scale in the lower left.

These values 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 dependences that may indicate systematic errors.

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

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 Section 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 binned the shear spatially with respect to the field of view.

In Fig. 15, we show the mean shear as a function of position on the focal plane for both IM3SHAPE (left) and NGMIX (right). Each whisker represents the mean shear of all galaxies that were ever observed in that area of the focal plane. As our shear measurements used information from multiple epochs, each measurement contributed to this plot multiple times: once for each single-epoch observation of that galaxy.

This figure is similar to Fig. 9, in which we showed the residual PSF pattern as a function of position on the focal plane. These plots are 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 catalogue. This is not surprising; we expected these PSF interpolation errors to leak into the galaxy shapes.

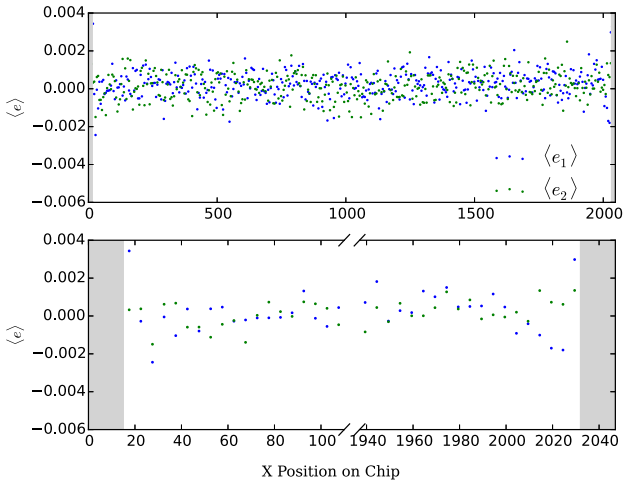

Figure 16. The mean shear $\langle e_1 \rangle$ (blue) and $\langle e_2 \rangle$ (green) for NGMIX, binned by column number X; X runs along the readout rows of the CCDs. (The corresponding plot for IM3SHAPE looks similar but noisier.) The bottom panel shows the same data blown up near the left and right 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.

8.1.2 Position on CCD

If we bin the shears by their column number, irrespective of the CCD number, as shown in Fig. 16, 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 fig. 6) showed that this effect led

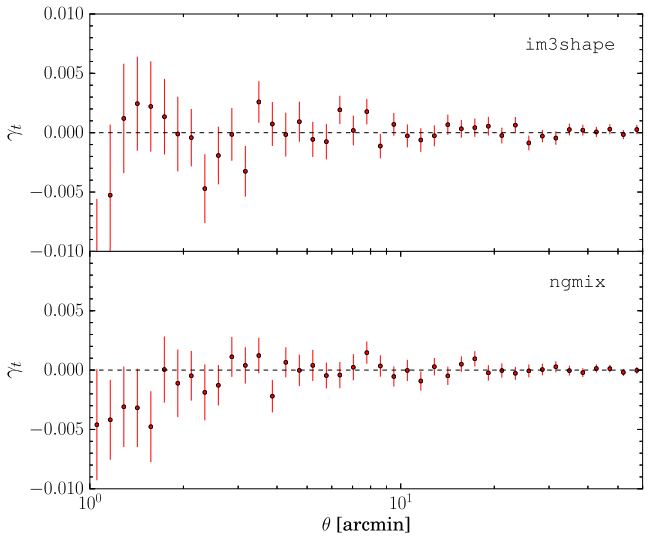

Figure 17. The tangential shear of galaxies in IM3SHAPE (top) and NGMIX (bottom) around field centres. Both measurements are approximately consistent with zero, although at scales less than 10 arcmin, both show a slight departure from the expected null signal. The magnitude of this effect is well below our requirements in both cases.

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.

In Fig. 16, we show 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 Fig. 16 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 estimated 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 per cent 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 (12), $c_{\rm rms} < 2 \times 10^{-3}$; however, it is not below the expected requirements for DES 5-yr 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 centres

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

In Fig. 17, we show 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 subfields. At large scales, the measurements are consistent with zero, but at scales less than about 10 arcmin there are a few consecutive bins with ∼1σ 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 is 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 misestimation of the distortion to be largest at a separation of around 1°. 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 centre, indicates that this is probably not due to errors in the WCS solution, although we do not have a good hypothesis for a plausible cause.

We estimated 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 $(2\,\text{arcmin}/62\,\text{arcmin})^2 = 0.001$. 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 was even smaller than the shear around the field centres.

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 looked 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.

8.2.1 PSF leakage

As we introduced in equation (5), we assumed 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:

$$e_{\rm gal} = e_{\rm true} + \alpha e_{\rm PSF} + c. \quad (43)$$

A measured slope of galaxy ellipticity versus 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 Fig. 18 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 per cent ± 3 per cent for IM3SHAPE and from −2.0 to 0.5 per cent ± 1 per cent for NGMIX. The slopes are consistent with no PSF leakage for both catalogues.

To obtain a more precise estimate of α, we computed the (weighted) average of the slopes of the red lines on the left plots (i.e. $\langle e_1 \rangle$ versus PSF e_1) and the blue lines on the right plots (i.e. $\langle e_2 \rangle$ versus PSF e_2). For IM3SHAPE, we found $\alpha = 0.008 \pm 0.025$, 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 16) given the uncertainty in the estimate.

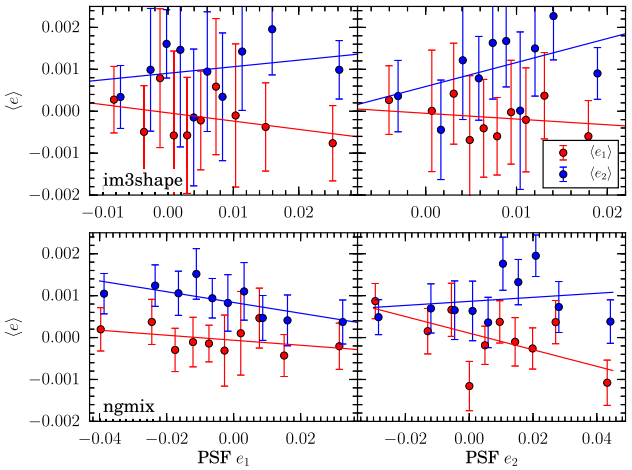

Figure 18. The mean galaxy shear as a function of the input PSF ellipticity (e_1 left, and e_2 right) for IM3SHAPE (top) and NGMIX (bottom). The solid lines show the best linear fit without binning. Note the range of the abscissa is different for the NGMIX and IM3SHAPE plots. The NGMIX measurements are 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 catalogues.

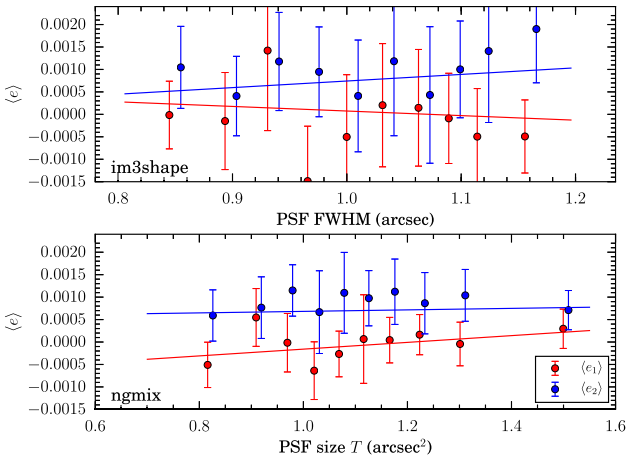

Figure 19. The mean galaxy shear as a function of the input PSF size for IM3SHAPE (top) and NGMIX (bottom). The solid lines show the best linear fit without binning.

We similarly plot the mean shear as a function of PSF size in Fig. 19 for both IM3SHAPE (left) and NGMIX (right). Linear best-fitting lines are also included. The slopes here are also consistent with zero, being on the order of 0.1 per cent 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 (43) and solving for α, we find that

$$\alpha = \frac{\xi_+^{gp} - \langle e_{\rm gal}\rangle^* \langle e_{\rm PSF}\rangle}{\xi_+^{pp} - |\langle e_{\rm PSF}\rangle|^2}, \tag{44}$$

where ξ_+^{pp} is the autocorrelation 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 is a quantification

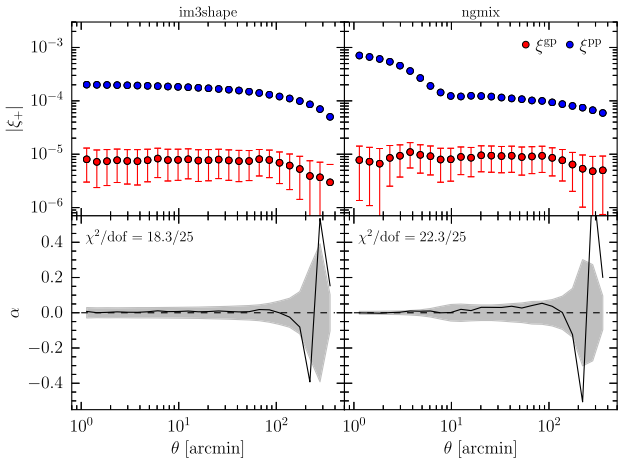

Figure 20. The calculation of the PSF leakage parameter α, which is given in equation (44). The top plots show ξ^{gp} (red), the cross-correlation of the galaxy shapes with the PSF shapes, and ξ^{pp} (blue), the autocorrelation of the PSF shapes, for IM3SHAPE (left) and NGMIX (right). The bottom plots 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 χ^2/d.o.f. is given for α over all scales.

of 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 Fig. 20 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 used the mock catalogues described in Becker et al. (2015) to compute the total uncertainty for α. These catalogues were populated with PSF shapes by using the PSF shape from the nearest observed galaxy to each mock galaxy. We then used the full suite of 126 mock catalogues to compute the total uncertainty on α including both shape noise and sample variance.

We found that both IM3SHAPE and NGMIX show no significant PSF contamination in this test, with a total χ^2/d.o.f. of 18.3/25 and 22.3/25 for α computed over all scales. The best-fitting 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 (16), although in the case of IM3SHAPE we are only able to constrain $|\alpha|$ to be less than 0.03 at about 1σ.

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 measured 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 we found the signal, shown in Fig. 21, to be 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 modelling. The bright stars are not themselves used to constrain the PSF model (cf. Fig. 6), so these stars instead check for problems related to deblending and sky

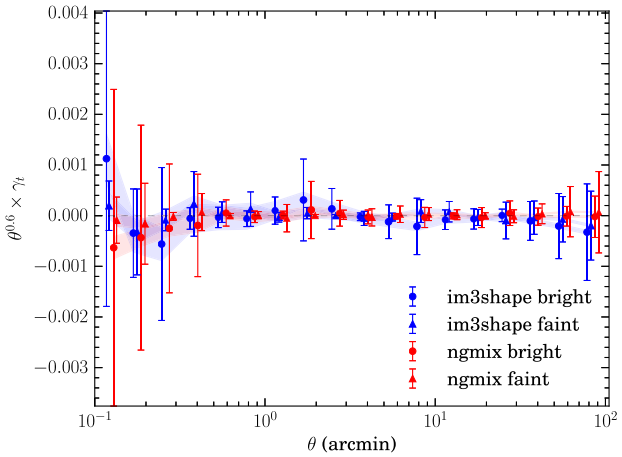

Figure 21. Tangential shear around stars for IM3SHAPE (blue) and NGMIX (red). Stars are split into to two bins of *i*-band magnitude, where 'bright' means $14 < m_i < 18.3$ and 'faint' means $18.3 < m_i < 22$. The faint sample includes stars used for PSF modelling; bright stars are excluded to avoid the brighter fatter effect (cf. Section 4.2). Shaded regions represent 1σ shape noise uncertainty, while error bars are from jackknifing the stars.

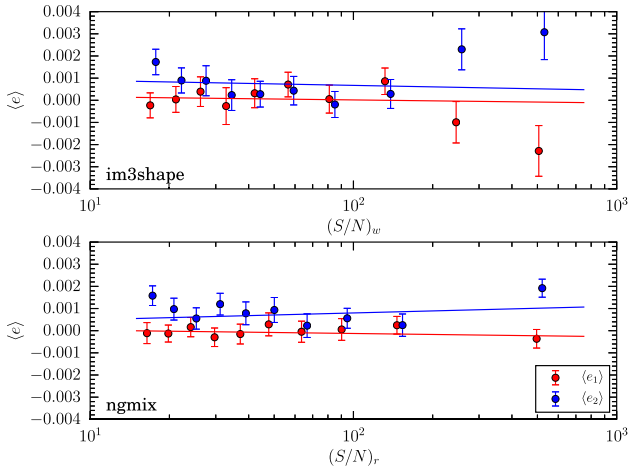

Figure 22. The mean galaxy shear as a function of the *S/N* for IM3SHAPE (top) and NGMIX (bottom). For IM3SHAPE, we test against $(S/N)_w$ (cf. equation 31), which is one of the parameters used for the calibration, so it includes corrections for selection bias. For NGMIX, we test against $(S/N)_r$ (cf. equation 32), which does not induce any significant selection bias from the binning.

estimation errors in the outskirts of bright stellar haloes. 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 neighbours being masked, the fraction of the stamp area being masked, the estimated bulge-to-disc ratio, the galaxy's *S/N*, 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 Section 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 S/N

The null test for checking that the galaxy shapes are independent of *S/N* requires different measures of *S/N* for each catalogue. As described in Section 7.3.2, for IM3SHAPE we calibrated the bias in the shear measurements 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. Section 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 Fig. 22. There is no apparent bias in the mean shear down to $(S/N)_w = 15$.

We did not apply any external calibration to NGMIX, so the null test for it requires an *S/N* measure that does not induce selection biases from the binning. For NGMIX, we used $(S/N)_r$ (cf. equation 32), which did not induce any selection biases when we tested it on simulated data. In the bottom panel of Fig. 22, we show that the mean estimated shear for NGMIX is independent of this 'roundified' *S/N* measure down to $(S/N)_r = 15$.

Previous versions of the NGMIX catalogue 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. In Fig. 22, we show that the final catalogues 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 12).

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. In the top panel of Fig. 23, we show 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[19] shear g, then the measured size will be

$$T^{(g)} = T^{(g=0)} \left(\frac{1 + |g|^2}{1 - |g|^2} \right). \quad (45)$$

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 antialigned. This can lead to an apparent

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

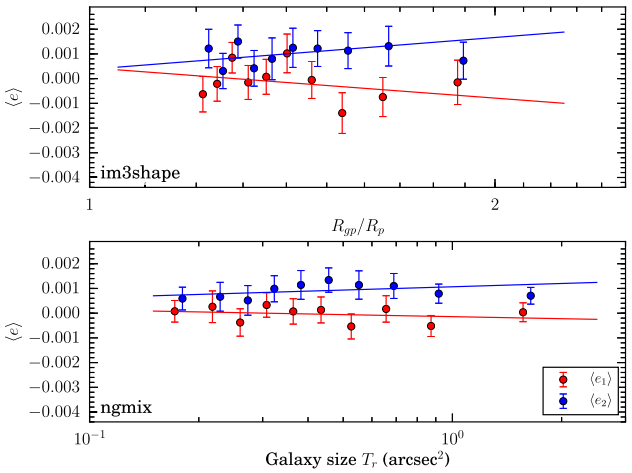

Figure 23. The mean galaxy shear as a function of the galaxy size for IM3SHAPE (top) and NGMIX (bottom). For IM3SHAPE, we test against R_{gp}/R_p, which is one of the parameters used for the calibration, so it includes corrections for selection bias. For NGMIX, we test against a size measure, T_r (cf. equation 46), that does not induce and significant selection bias from the binning.

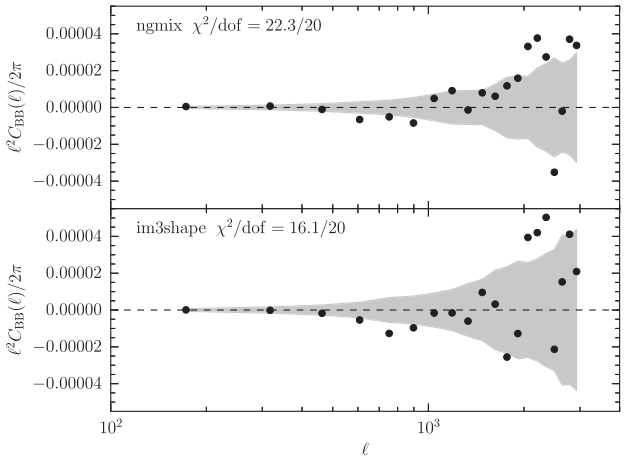

Figure 24. The measured B-mode for NGMIX (top) and IM3SHAPE (bottom). Each band power measurement is plotted at its central location in ℓ. The grey band shows the uncertainty on the measurement due to both sample variance and shape noise. Adjacent points are highly correlated and the indicated χ^2 accounts for the correlations.

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 *S/N* 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:

$$T_r \equiv T \left(\frac{1 - |e|^2}{1 + |e|^2} \right), \tag{46}$$

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. In the lower panel of Fig. 23, we show 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.

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 12).

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, van Waerbeke & Mellier 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 calculated 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 chose linear combinations that approximate band-powers in Fourier space as described in Becker & Rozo (2016). Finally, we used the mock catalogues described in Becker et al. (2015) to compute the shape noise and sample variance uncertainty for the statistics. These mock catalogues include the survey mask and match the shape noise and source photometric redshift distribution for each of the two shear catalogues. We used 126 mock catalogues in total.

In Fig. 24, we show the measured B-mode for each catalogue 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 χ^2/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 Section 6.1 to test the performance of the two shear algorithms on relatively realistic images with known applied shear.

Since IM3SHAPE uses this simulation to calibrate the shear measurements (cf. Section 7.3.2), the overall corrected shears should be accurate, almost by construction. The calibration was done without weighting, but for this test we used the same weights that we recommend for the data (cf. Section 7.3.4). The mean shear is thus not mathematically guaranteed to be exactly zero. We detect a net bias after applying the calibrations, but it is less than our requirements for DES SV data: $m_1 = 0.0008 \pm 0.0015$ 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. Section 6.1). We found the overall calibration error for NGMIX to be $m_1 = -0.030 \pm 0.0015$ and $m_2 = -0.035 \pm 0.0015$. This does not quite meet our requirement of $|m| < 0.03$ from equation (11).

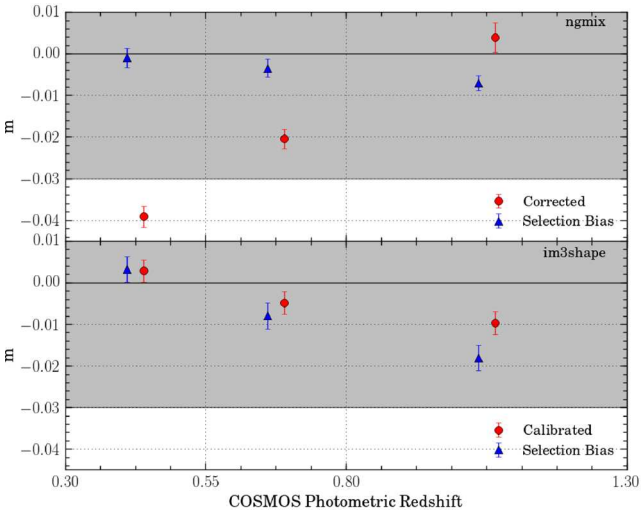

Figure 25. Multiplicative shear bias in three bins of photometric redshift for NGMIX (top) and IM3SHAPE (bottom), as calculated using the GREAT-DES simulation. In both cases, the same selection and weights were used as for the real data. The red circles denote the average bias in each bin after correcting 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 ±3 per cent requirement for SV data.

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. To test the redshift dependence of the bias, we used the known photometric redshifts of the galaxies from the COSMOS data, from which we drew the galaxy images used in the GREAT-DES simulation. With this test, we can also learn if the tomographic selection process itself leads to any significant selection biases.

In Fig. 25, we show the results of performing this test for NGMIX (top) and IM3SHAPE (bottom), taking galaxies in different ranges of photometric redshift, using the same 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, Fig. 25), 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). We found the overall leakage before calibration was $\alpha = (0.070, 0.112, 0.102)$ for the three redshift bins. After calibration, we found $\alpha = (0.001, 0.021, -0.005)$, which demonstrates good performance of the leakage calibration as well.

We found the multiplicative bias for NGMIX to be outside of the requirement band for the lowest redshift bin, and then rose 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 disc model, we implemented a more flexible model and applied it to this simulation. This model is a simple two-component bulge and disc 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 disc 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 found biases ∼1 per cent, 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 Fig. 25 represent the estimated selection bias in each bin induced by our various selection criteria. To calculate these values, we applied the known shear to the COSMOS shape estimates (Kannawadi et al. 2015) of the galaxies used for the simulation, and then applied the same selection criteria we used for each of the two algorithms. We found the selection bias from the IM3SHAPE cuts was at most 2 per cent in the highest redshift bin, which is largely corrected by the calibration scheme. We found the selection bias for NGMIX was less than 1 per cent for all redshift bins.

8.6 Cross-catalogue comparisons

Another powerful test is to compare the two independent shear catalogues, IM3SHAPE and NGMIX. We used two very different strategies when generating these catalogues. For IM3SHAPE, we used simulations to determine the shear calibration, and applied corrections to the shear measurements on real data. For NGMIX, we expected relatively little noise bias, but the sensitivity of the shear estimator was calculated from the data itself and applied to the shear measurements, a process that was worth testing in detail. And we did expect some model bias for NGMIX. Furthermore, the PSF was treated differently by the two methods: for IM3SHAPE, we used 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-catalogue comparison, since there is not a unique unbiased shear estimate for a single galaxy. Rather, we wished to test that both methods produced 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 tested if the results were consistent when using the same ensemble of galaxies with the same weighting.

Disagreement between the catalogues would be proof that at least one catalogue is biased, but we would not be able to determine which one, nor the magnitude of this bias. Agreement between the two catalogues is subjectively reassuring, but we wish to emphasize that agreement does not prove that both catalogues 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 catalogues, 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 catalogues.

For this test, we used the tangential shear signal around luminous red galaxies (LRGs) as determined by redMaGiC (red sequence Matched-filter Galaxies Catalogue; Rozo et al. 2015) from the same DES SPT-E data. For this purpose, we did not require sources to be behind the lenses. Rather, we took the full LRG catalogue as

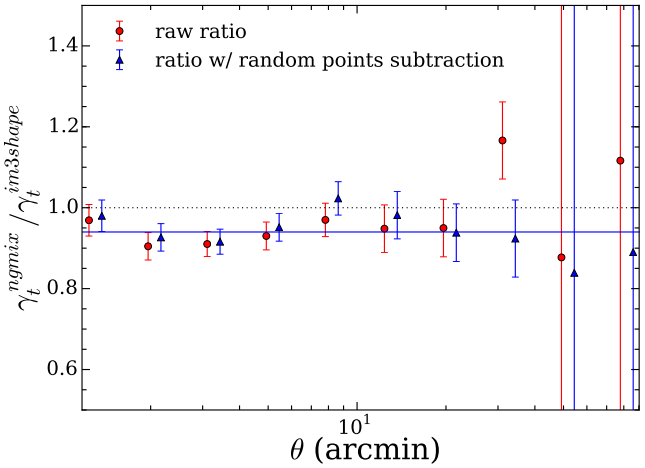

Figure 26. The ratios of tangential shear measurements around LRG galaxies from shears measured by NGMIX to those measured by IM3SHAPE. The red circles show the direct ratio and the triangles correspond to the ratio after subtraction of the tangential shear around random points. The weighted mean ratio in the scale range 1–20 arcmin 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 catalogues. This result is in good agreement with the data points.

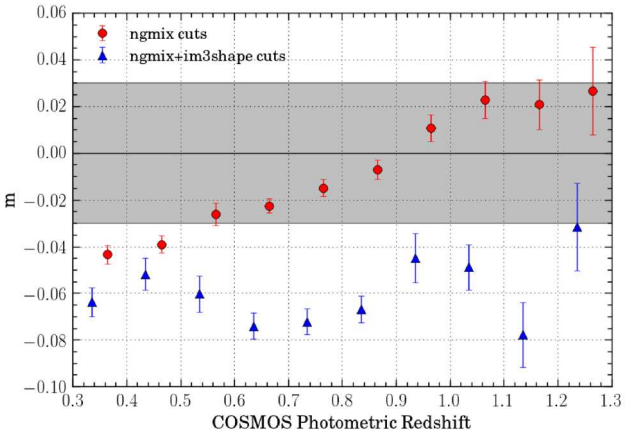

Figure 27. Multiplicative bias for NGMIX shear measurements on GREAT-DES simulated data as a function of redshift. The red circles show the bias calculated using all galaxies that pass the NGMIX selection criteria (as in the upper panel of in Fig. 25). The blue triangles show the bias when also including the recommended IM3SHAPE selection, as we do to obtain the matched catalogue used for Fig. 26. As in Fig. 25, the grey band represents the ±3 per cent requirement for the SV data.

the lenses, and for the sources, we used all galaxies that were well measured by both NGMIX and IM3SHAPE. Regardless of the redshifts of the LRGs and the source galaxies, we expected the signal to be the same for both catalogues in the absence of a multiplicative bias.

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

$$\langle e_{t,i}(\theta)\rangle = (1+m_i)\langle\gamma_t(\theta)\rangle + \langle\eta_i(\theta)\rangle, \quad (47)$$

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 catalogues have identical values 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 Fig. 26 represent the ratio of measured tangential shear using the two shear catalogues. The weighted mean of the ratio over the range from 1 to 20 arcmin (the typical scales of interest for weak lensing) is 0.932 ± 0.018. We would naively expect this to be an estimate of $(1+m_{\rm NGMIX})/(1+m_{\rm IM3SHAPE}) \approx 1+m_{\rm NGMIX}-m_{\rm IM3SHAPE}$. 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 Fig. 26 represent the signal after this subtraction, and have a mean ratio of 0.954 ± 0.018.

Secondly, 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 created simulated realizations of the ratio, and compared the measured signal to the mean and variance of these. We generated a ratio realization in the following way.

(i) 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, $\hat{\gamma}_t(\theta)$.

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

(iii) 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.

(iv) Compute the realization ratio as

$$\left(\hat{\gamma}_t + \langle\eta^r_{\rm NGMIX}\rangle\right) / \left(\hat{\gamma}_t + \langle\eta^r_{\rm IM3SHAPE}\rangle\right). \quad (48)$$

We found the mean of these realizations to be consistent with a ratio of 1 on all scales, which means that the S/N of the tangential shear is high enough that we can neglect the noise term in the denominator.

Finally, we found that the act of matching the two catalogues caused a selection bias in the NGMIX catalogue, for two reasons. First, the IM3SHAPE algorithm failed more often for objects with low Sérsic index ($n < 1$). And secondly, the cuts we made on the IM3SHAPE measurements of $(S/N)_w$ and R_{gp}/R_p altered the mix of galaxy properties in the matched catalogue. These two selection effects, when applied to the NGMIX catalogue imparted a net bias on the NGMIX shear estimates in the matched catalogue that was not present in the full NGMIX catalogue.

We quantified the level of this selection bias by performing the same procedure on the GREAT-DES simulation. We compared the mean bias for NGMIX using the canonical NGMIX selection criteria to the bias after applying the IM3SHAPE selection, as a function of redshift. The result is shown in Fig. 27. The matching induced a mean selection bias of about −3 per cent. Furthermore, we found that this bias increased 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 found a net selection bias of −6 per cent for NGMIX in the matched catalogue relative to whatever bias might be present in the full NGMIX catalogue.[20]

The mean ratio of 0.954 ± 0.018 is thus consistent with the prediction from GREAT-DES of −6 per cent 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 catalogues is shown by the blue line in Fig. 26. Our finding is therefore consistent with no relative multiplicative bias between the two catalogues.

We cannot of course prove that neither catalogue 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 uncertainties in the calculation of the predicted selection bias described above that may be at the ∼3 per cent 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 catalogues:

$$\xi_{+,\Delta e}(\theta) = \langle (e_{\rm NGMIX}(\boldsymbol{x}) - e_{\rm IM3SHAPE}(\boldsymbol{x}))^* \\ (e_{\rm NGMIX}(\boldsymbol{x}+\boldsymbol{\theta}) - e_{\rm IM3SHAPE}(\boldsymbol{x}+\boldsymbol{\theta}))\rangle. \quad (49)$$

Consider the following model for the additive systematic errors in each catalogue (labelled i here):

$$e_i = (1+m_i)\gamma + \eta_i + a_i c_{\rm common} + c_i, \quad (50)$$

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 catalogues, possibly multiplied by different coefficients a_i, and c_i is the additive error particular to each catalogue.

By construction, the additive bias terms in equation (50) 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

$$\xi_{+,\Delta e}(\theta) = (\Delta m)^2 \xi_+(\theta) \\ + (\Delta a)^2 \langle c^*_{\rm common} c_{\rm common}\rangle(\theta) \\ + \langle c^*_{\rm NGMIX} c_{\rm NGMIX}\rangle(\theta) \\ + \langle c^*_{\rm IM3SHAPE} c_{\rm IM3SHAPE}\rangle(\theta). \quad (51)$$

This test is sensitive to the spatial correlations of the systematic errors in either catalogue, 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 Section 8.6.1, the act of matching the two catalogues can induce selection biases that are not present in either catalogue 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.

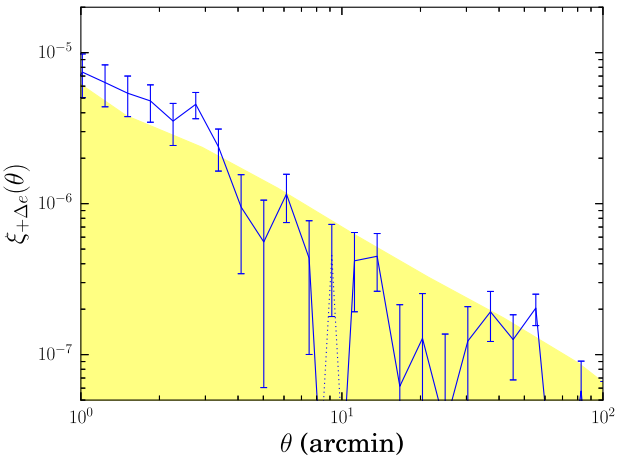

Figure 28. The shear autocorrelation function of the difference in shear estimates of NGMIX and IM3SHAPE. This test shows the level of additive systematic errors that may still be present in one catalogue that is not present in the other. The yellow band is the requirement, $\delta\xi_+^{\rm max}$ from Fig. 3.

The estimated value of α for NGMIX changed by less than 0.1 per cent in the matched catalogue relative to the full NGMIX catalogue. But for IM3SHAPE, the matching changed α by −1.5 per cent. Therefore, to make this a fair test of the additive systematic errors, we added back $0.015 \times e_{\rm PSF}$ to the IM3SHAPE galaxy shapes to account for this selection effect.[21]

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

In Fig. 28, we show the resulting correlation function (equation 49) after subtracting these selection biases. For the weights, we used $w = \sqrt{w_{\rm NGMIX} \times w_{\rm IM3SHAPE}}$. The yellow band represents our requirement for additive systematic errors from equation (13). We see that, at scales less than 3 arcmin, we do not quite meet the requirements. Either one or both catalogues apparently have non-negligible additive systematic errors at these scales. We recommend that science applications sensitive to additive systematic errors check carefully 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 catalogues, given all of the information we have available.

In Table 2, we provide a summary of the results from the previous sections (including the tests in Section 4). For each, we have

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

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

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

Test	Upper limit on systematic error	
	IM3SHAPE	NGMIX
PSF model tests		
Section 4.2 Mean PSF size error	$\|m\| < 0.005$	$\|m\| < 0.01$
Section 4.4 PSF model diagnostics $\rho_{1,3,4}$	$\xi_+^{cc}(1\,\text{arcmin}) < 2 \times 10^{-6}$ [a] $\xi_+^{cc}(30\,\text{arcmin}) < 7 \times 10^{-8}$	$\xi_+^{cc}(1\,\text{arcmin}) < 5 \times 10^{-6}$ $\xi_+^{cc}(30\,\text{arcmin}) < 9 \times 10^{-8}$
Section 4.4 PSF model diagnostics $\rho_{2,5}$	$\xi_+^{cc}(1\,\text{arcmin}) < 2 \times 10^{-7}$ $\xi_+^{cc}(30\,\text{arcmin}) < 1.5 \times 10^{-7}$	$\xi_+^{cc}(1\,\text{arcmin}) < 8 \times 10^{-8}$ $\xi_+^{cc}(30\,\text{arcmin}) < 1.4 \times 10^{-7}$
Spatial tests		
Section 8.1.1 Position in the field of view	No evidence of systematic errors	No evidence of systematic errors
Section 8.1.2 Position on CCD	$\xi_+^{cc}(1\,\text{arcmin}) < 6 \times 10^{-7}$	$\xi_+^{cc}(1\,\text{arcmin}) < 6 \times 10^{-7}$
Section 8.1.3 Tangential shear around field centres	$\xi_+^{cc}(1\,\text{arcmin}) < 4 \times 10^{-8}$	$\xi_+^{cc}(1\,\text{arcmin}) < 4 \times 10^{-8}$
PSF tests		
Section 8.2.1 PSF leakage	$\|\alpha\| < 0.04$	$\|\alpha\| < 0.01$
Section 8.2.1 Dependence on PSF size	No evidence of systematic errors	No evidence of systematic errors
Section 8.2.2 Star–galaxy cross-correlation	$\|\alpha\| < 0.03$	$\|\alpha\| < 0.015$
Section 8.2.3 Tangential shear around stars	No evidence of systematic errors	No evidence of systematic errors
Galaxy property tests		
Section 8.3.1 Galaxy S/N	No evidence of systematic errors	No evidence of systematic errors
Section 8.3.2 Galaxy size	No evidence of systematic errors	No evidence of systematic errors
B-mode statistics		
Section 8.4 $\ell^2 C_{\rm BB}(\ell)/2\pi$	No evidence of B-mode	No evidence of B-mode
Calibration tests		
Section 8.5 Redshift dependence in GREAT-DES	$\|m\| < 0.02$	$\|m\| < 0.04$
Cross-catalogue comparison		
Section 8.6.1 Tangential shear ratio	$\|\Delta m\| \lesssim 0.04$	
Section 8.6.2 Differential shear correlations	$\|\xi_+^{cc}(1\,\text{arcmin})\| < 9 \times 10^{-6}$ $\|\xi_+^{cc}(30\,\text{arcmin})\| < 2 \times 10^{-7}$	

[a]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 Fig. 10, which uses *r*, *i*, *z* bands.

converted the result of the test into the impact that the result could have on four 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 $\theta = 1$ arcmin and (when relevant) 30 arcmin. 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 arcmin and 30 arcmin 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 arcmin 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 modelling systematic errors act as $c_{\rm common}$ terms in the nomenclature of Section 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 catalogue:

$$\text{IM3SHAPE} \quad \begin{matrix} \left|\xi_+^{\rm sys}(1\,\text{arcmin})\right| < 1.1 \times 10^{-5} \\ \left|\xi_+^{\rm sys}(30\,\text{arcmin})\right| < 4 \times 10^{-7} \end{matrix} \tag{52}$$

$$\text{NGMIX} \quad \begin{matrix} \left|\xi_+^{\rm sys}(1\,\text{arcmin})\right| < 1.4 \times 10^{-5} \\ \left|\xi_+^{\rm sys}(30\,\text{arcmin})\right| < 4 \times 10^{-7}. \end{matrix} \tag{53}$$

Note that we are not claiming that either catalogue 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 catalogues 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 catalogues 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 catalogues.

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, Fig. 12 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 catalogues 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 catalogues is

$$|m| < 0.05. \quad (54)$$

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

9 SHEAR CATALOGUES

The final shear catalogues are publicly available on the DESDM Releases web page.[22] See that web page for complete documentation about how to access these catalogues, as well as the other DES SV catalogues that are available.

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 B has further details about the content and structure of the catalogues.

9.1 Final galaxy selection

The starting point for our galaxy catalogues was described in Sections 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 describe further cuts informed by the suite of null tests in Section 8, such that the final shear catalogues were found to pass our tests.

We removed individual objects according to the following criteria.

(i) SExtractor flags = 1 or 2. Objects with higher SExtractor flags were already been removed from the input catalogues, since they are clearly problematic for a shape measurement code. But these two flags indicate that the object is likely to be blended, and thus the shape measurement was likely to be corrupted.

(ii) 'Crazy colours'.[23] Individual objects with questionable colours are probably contaminated by cosmic rays or other defects, so their shapes are also likely to be inaccurate.

(iii) Very low surface brightness. We found a class of spurious objects with very large sizes, but relatively low flux that were usually associated with various image artefacts. We excluded 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. This cut is efficient at bright fluxes, but less so at faint fluxes, removing a significant number of real galaxies.

(iv) Tiny size. If the NGMIX estimate of the object size is very small, then the object is probably a star. Specifically, we removed objects with $T + \sigma_T < 0.02$ square arcsec.

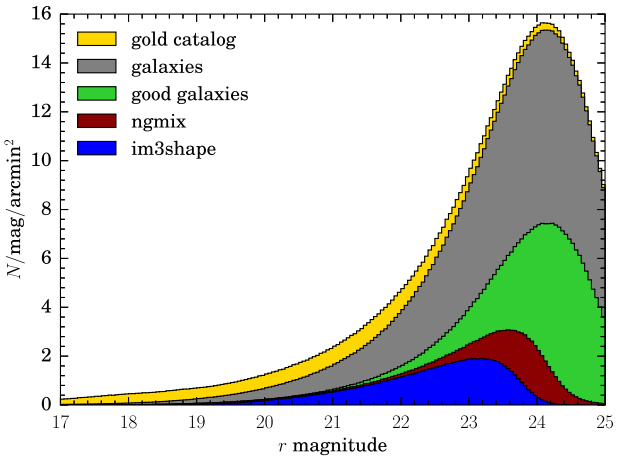

Figure 29. A histogram of the r-band magnitude distribution showing the application of the various selection criteria from the initial 'Gold' Catalogue to the two final shear catalogues. The dark red and blue show the galaxies with sufficiently accurate shape for NGMIX and IM3SHAPE, respectively.

From the resulting set of 'good galaxies', we then made 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 passed the null tests.

As we have already mentioned in Section 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 32 and 46) to avoid inducing a selection bias.

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

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

We used this selection for all of the test results shown in Section 8.

In Fig. 29, we show the effect that successively applying each round of selections has on the distribution of r-band magnitudes, starting with the original Gold Catalogue, selecting possible galaxies, removing problematic galaxies, and then applying the S/N and size criteria for the two shear catalogues.

9.2 Applying the calibration/sensitivity

For both IM3SHAPE and NGMIX, the raw galaxy shape values given in the catalogue are intrinsically biased estimators of the shear. In the case of IM3SHAPE, simulation-based calibration is used (cf. Section 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 was calculated and given in the catalogue (cf. Section 7.4.3).

[22] http://des.ncsa.illinois.edu/releases/sva1

[23] 'Crazy colours' mean any of the following: $g - r < -1$, $g - r > 4$, $i - z < -1$, or $i - z > 4$.

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:

$$\langle\gamma\rangle = \frac{\sum(e_i - c_i)}{\sum s_i}, \tag{55}$$

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 catalogue comes with a recommended weight w_i to use for making these ensemble averages:

$$\langle\gamma\rangle = \frac{\sum w_i(e_i - c_i)}{\sum w_i s_i}. \tag{56}$$

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

$$\langle\gamma^a\gamma^b\rangle = \frac{\sum w_i^a w_j^b (e_i^a - c_i^a)(e_j^b - c_j^b)}{\sum w_i^a w_j^b s_i^a s_j^b}. \tag{57}$$

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):

$$\mathrm{var}(\langle\gamma_{1,2}\rangle) \equiv \frac{\sigma_{\rm SN}^2}{\Omega n_{\rm eff}}, \tag{58}$$

where $\sigma_{\rm SN}$ is the shape noise per component.

Applying all of the selections defined in Section 9.1 to our shape catalogues 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. Section 2.1), which leads to direct number densities of 4.2 and 6.9 galaxies per square arcmin, respectively.

To turn these numbers into proper effective number densities, we first need to calculate the shape noise $\sigma_{\rm SN}^2$,

$$\sigma_{\rm SN}^2 = \frac{\sum w_i^2\left(|e_i|^2 - 2\sigma_{e,i}^2\right)}{2\sum w_i^2 s_i^2}, \tag{59}$$

where $2\sigma_e^2$ is the trace of the covariance matrix of e_1, e_2,[24] 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 catalogues.

[24] 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 $1/(\sigma_{\rm SN}^2 + \sigma_e^2)$, cf. Section 7.3.4.

For the IM3SHAPE catalogue, this number comes to $\sigma_{\rm SN} = 0.233$, and for NGMIX, $\sigma_{\rm SN} = 0.243$.

The variance of each component of the mean shear over the entire survey area can be calculated from equation (56):

$$\mathrm{var}(\langle\gamma_{1,2}\rangle) = \frac{\sum w_i^2\left(s_i^2\sigma_{\rm SN}^2 + \sigma_{e,i}^2\right)}{\left(\sum w_i s_i\right)^2}, \tag{60}$$

which, using equations (58) and (59), leads to

$$n_{\rm eff} = \frac{1}{\Omega}\frac{\sigma_{\rm SN}^2\left(\sum w_i s_i\right)^2}{\sum w_i^2\left(s_i^2\sigma_{\rm SN}^2 + \sigma_{e,i}^2\right)} \tag{61}$$

$$= \frac{1}{\Omega}\frac{\left(\sum w_i s_i\right)^2}{\sum w_i^2 s_i^2}\left(1 - \frac{2\sum w_i^2\sigma_{e,i}^2}{\sum w_i^2|e_i|^2}\right). \tag{62}$$

For IM3SHAPE, we find $n_{\rm eff} = 3.7$ galaxies per square arcmin, and for NGMIX, $n_{\rm eff} = 5.7$ galaxies per square arcmin.

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

$$n_{\rm eff} = \frac{1}{\Omega}\frac{\left(\sum w_i\right)^2}{\sum w_i^2}. \tag{63}$$

Using this definition, we obtain $n_{\rm eff} = 4.1$ and 6.8 for IM3SHAPE and NGMIX, respectively. With this definition however, the appropriate numerator in the ratio $\sigma_\epsilon^2/n_{\rm eff}$ is not the intrinsic shape noise $\sigma_{\rm SN}^2$, 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 arcmin that was predicted for DES (The Dark Energy Survey Collaboration 2005). This is in part because of our decision to cut both catalogues 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 catalogue and 1.0 million from the NGMIX catalogue. 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 catalogues 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 5 yr. Instead, the mean is approximately seven exposures averaged across the SPT-E area. If we reach an average of 10 exposures, this will lead to a 20 per cent 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 arcsec, while the median seeing during SV was slightly above 1.0 arcsec. We are closer to achieving our goal of 0.9 arcsec 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 neighbours. The SEXTRACTOR flags related to blended objects removed almost 1 million galaxies from the catalogues. We are currently working on an algorithm to model the profiles of neighbouring 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 catalogue.

Another obvious improvement will be to use multiband 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 artefacts 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 artefacts, removed 1.5 million objects. Presumably many of these are real galaxies rather than image artefacts, so if we can omit this step, we will keep more galaxies in the catalogue.

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

10 SUMMARY AND DISCUSSION

We present here two shear catalogues for the SPT-E region observed as part of the DES SV time. Both catalogues, 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 catalogues have 4.2 and 6.9 galaxies per square arcmin (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 arcmin, respectively (cf. Section 9.3).

For creation of both shear catalogues, we used 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).

In addition to passing null tests on the data individually, the two catalogues are 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 catalogues 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 catalogues, and thus the first to be able to show this kind of consistency.

In Section 8.7, we estimated upper limits on the level of additive systematic errors that may be present in the two catalogues at 1 arcmin and 30 arcmin. We recommend a Bayesian prior of $|m| < 0.05$ for the systematic uncertainty on the calibration for both catalogues. Since NGMIX is the deeper catalogue (due to using multi-band data rather than just r band), users desiring the most precise measurement may adopt NGMIX as their canonical shear catalogue. However, we strongly recommend also comparing with results using the IM3SHAPE catalogue, while carefully taking into account any selection effects; any discrepancy indicates a systematic error in one or both catalogues.

While the catalogues were seen to be sufficiently accurate for SV weak lensing science, they do not yet pass the tests at the level that will be needed for the full 5-yr 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 modelling (cf. Figs 6 and 10). Fortunately, there has been a significant amount of work in recent years on improved PSF modelling and interpolation algorithms (e.g. Chang et al. 2012; Li, Xin & Cui 2012; Gentile, Courbin & Meylan 2013; Kitching 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, Reil & Davis 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. Fig. 16). We expect this to reduce some of the spurious features seen in Figs 9 and 17.

We have recently implemented an algorithmic correction to the brighter fatter relation discussed in Section 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 neighbours 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 catalogues. Contamination by neighbours was one of the more significant cuts that led to the drop in number density for the 'good galaxies' seen in Fig. 29.

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, a preliminary version of which was tested in the GREAT3 challenge (Mandelbaum et al. 2015). Considering how useful we found it to have two catalogues, we are looking forward to the prospect of additional catalogues to compare in various ways.

We also plan to begin 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-yr 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 SV 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 catalogues in future DES analyses. We must keep the level of systematic errors below the improved statistical uncertainty for these data. The full 5-yr 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 SV 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.

We also thank the anonymous referee, whose careful reading helped to significantly improve the paper.

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 US Department of Energy, the US 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 DES.

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, LBNL, 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.

REFERENCES

Abazajian K. 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, J. Instrum., 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., 2016, MNRAS, 457, 304
Becker M. R. et al., 2015, preprint (arXiv:1507.05598)
Bergé J., Gamper L., Réfrégier A., Amara A., 2013, Astron. Comput., 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 Gabriel C., Arviset C., Ponz D., Enrique S., eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac., San Francisco, p. 112
Bertin E., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser. Vol. 442, Astronomical Data Analysis Software and Systems XX. Astron. Soc. Pac., San Francisco, p. 435
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, ASP Conf. Ser. Vol. 281, Astronomical Data Analysis Software and Systems XI. Astron. Soc. Pac., San Francisco, p. 228
Bonnett C., 2015, MNRAS, 449, 1043
Bonnett C. et al., 2015, preprint (arXiv:1507.05909)
Bosch J. F., 2011, PhD thesis, Univ. California
Bouy H., Bertin E., Moraux E., Cuillandre J.-C., Bouvier J., Barrado D., Solano E., Bayo A., 2013, A&A, 554, A101
Bridle S. et al., 2009, Ann. Appl. Stat., 3, 6
Bridle S. et al., 2010, MNRAS, 405, 2044
Bruderer C., Chang C., Refregier A., Amara A., Berge J., Gamper L., 2016, ApJ, 817, 25
Carlstrom J. E. et al., 2011, PASP, 123, 568
Chang C. et al., 2012, MNRAS, 427, 2572
Chang C. et al., 2013, MNRAS, 434, 2121
Crittenden R. G., Natarajan P., Pen U.-L., Theuns T., 2001, ApJ, 559, 552
de Vaucouleurs G., 1948, Ann. Astrophys., 11, 247
Dempster A. P., Laird N. M., Rubin D. B., 1977, J. R. Stat. Soc., 39, 1
Desai S. et al., 2012, ApJ, 757, 83
Diehl T., 2012, Phys. Procedia, 37, 1332
Diehl H. T. et al., 2014, Proc. SPIE, 9149, 91490V
Duda R. O., Hart P. E., 1972, Commun. ACM, 15, 11
Flaugher B. et al., 2015, ApJ, 150, 150
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fu L. 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., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Gruen D. et al., 2013, MNRAS, 432, 1455
Gruen D., Bernstein G. M., Jarvis M., Rowe B., Vikram V., Plazas A. A., Seitz S., 2015, J. Instrum., 10, C05032
Guyonnet A., Astier P., Antilogus P., Regnault N., Doherty P., 2015, A&A, 575, A41
Heymans C. et al., 2005, MNRAS, 361, 160
Heymans C. et al., 2006, MNRAS, 368, 1323
Heymans C., Rowe B., Hoekstra H., Miller L., Erben T., Kitching T., van Waerbeke L., 2012a, MNRAS, 421, 381
Heymans C. et al., 2012b, MNRAS, 427, 146
Hirata C., Seljak U., 2003, MNRAS, 343, 459
Hirata C. M. et al., 2004, MNRAS, 353, 529
Hoekstra H., 2013, preprint (arXiv:1312.5981)
Hoffleit D., Jaschek C., 1991, in Hoffleit D., Jaschek C., eds, The Bright Star Catalogue, 5th rev. ed. Yale Univ. Observatory, New Haven, Conn.
Hogg D. W., Lang D., 2013, PASP, 125, 719
Honscheid K. et al., 2012, Proc. SPIE, 8451, 845112

Hough P. V. C., 1959, Proc. 2nd Int. Conf. High Energy Accelerators and Instrumentation, Vol. C590914, Machine Analysis Of Bubble Chamber Pictures, p. 554
Jarvis M., Bernstein G. M., Fischer P., Smith D., Jain B., Tyson J. A., Wittman D., 2003, AJ, 125, 1014
Jee M. J., Tyson J. A., Schneider M. D., Wittman D., Schmidt S., Hilbert S., 2013, ApJ, 765, 74
Kacprzak T., Zuntz J., Rowe B., Bridle S., Refregier A., Amara A., Voigt L., Hirsch M., 2012, MNRAS, 427, 2711
Kacprzak T., Bridle S., Rowe B., Voigt L., Zuntz J., Hirsch M., MacCrann N., 2014, MNRAS, 441, 2528
Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460
Kannawadi A., Mandelbaum R., Lackner C., 2015, MNRAS, 449, 3597
Kitching T. et al., 2011, Ann. Appl. Stat., 5, 2231
Kitching T. D. et al., 2012, MNRAS, 423, 3163
Kitching T. D. et al., 2013, ApJS, 205, 12
Klein J. R., Roodman A., 2005, Annu. Rev. Nucl. Part. Sci., 55, 141
Koekemoer A. M. et al., 2007, ApJS, 172, 196
Krause E., Hirata C. M., 2010, A&A, 523, A28
Kuijken K. et al., 2015, MNRAS, 454, 3500
Leauthaud A. et al., 2007, ApJS, 172, 219
Leistedt B. et al., 2015, preprint (arXiv:1507.05647)
Levenberg K., 1944, Q. Appl. Math., 2, 164
Lewis A., 2009, MNRAS, 398, 471
Li G., Xin B., Cui W., 2012, preprint (arXiv:1203.0571)
Lourakis M., 2004, levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++. Available at: http://www.ics.forth.gr/lourakis/levmar/.
Mandelbaum R. et al., 2005, MNRAS, 361, 1287
Mandelbaum R., Hirata C. M., Leauthaud A., Massey R. J., Rhodes J., 2012, MNRAS, 420, 1518
Mandelbaum R. et al., 2014, ApJS, 212, 5
Mandelbaum R. et al., 2015, MNRAS, 450, 2963
Marquardt D., 1963, SIAM J. Appl. Math., 11, 431
Massey R., Refregier A., 2005, MNRAS, 363, 197
Massey R. et al., 2007, MNRAS, 376, 13
Melchior P., Viola M., 2012, MNRAS, 424, 2757
Melchior P., Böhnert A., Lombardi M., Bartelmann M., 2010, A&A, 510, A75
Melchior P., Viola M., Schäfer B. M., Bartelmann M., 2011, MNRAS, 412, 1552
Melchior P. 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. et al., 2013, MNRAS, 429, 2858
Mohr J. J. et al., 2012, Proc. SPIE, 8451, 84510D
Nakajima R., Bernstein G., 2007, AJ, 133, 1763
Nelder J. A., Mead R., 1965, Comput. J., 7, 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, Proc. SPIE, 9145, 914516
Rowe B., 2010, MNRAS, 404, 350
Rowe B. T. P. et al., 2015, Astron. Comput., 10, 121
Rozo E. et al., 2015, preprint (arXiv:1507.05460)
Rykoff E. S. et al., 2016, ApJS, 224, 1
Sánchez C. et al., 2014, MNRAS, 445, 1482
Schneider P., van Waerbeke L., Mellier Y., 2002, A&A, 389, 729
Schrabback T. et al., 2007, A&A, 468, 823
Schrabback T. et al., 2010, A&A, 516, A63
Seitz C., Schneider P., 1997, A&A, 318, 687
Sérsic J. L., 1963, Bol. Asoc. Argentina Astron. La Plata Argentina, 6, 41
Sheldon E. S., 2014, MNRAS, 444, L25
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Soumagnac M. T. et al., 2015, MNRAS, 450, 666
Tarlé G. et al., 2010, Proc. SPIE, 7739, 77393L
The Dark Energy Survey Collaboration, 2005, preprint (astroph/0510346)
The Dark Energy Survey Collaboration et al., 2015, preprint (arXiv:1507.05552)
Tyson J. A., Roat C., Bosch J., Wittman D., 2008, in Argyle R. W., Bunclark P. S., Lewis J. R., eds, ASP Conf. Ser. Vol. 394, Astronomical Data Analysis Software and Systems XVII. Astron. Soc. Pac., San Francisco, p. 107
Van Waerbeke L. 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. 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., Hirsch M., Rowe B., Bridle S., 2013, MNRAS, 434, 1604

APPENDIX A: MULTI-EPOCH DATA STRUCTURES

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 catalogue, 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 three 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.

Table B1. Error flags in the IM3SHAPE catalogue. Objects with non-zero ERROR_FLAG should be removed from any science analysis.

Value	Decimal	Meaning
2^0	1	IM3SHAPE failed completely.
2^1	2	Minimizer failed to converge.
2^2	4	Tiny ellipticity $\|e\| < 10^{-4}$: IM3SHAPE fit failed.
2^3	8	e_1 or e_2 outside $(-1, 1)$.
2^4	16	Radius >20 arcsec.
2^5	32	$R_{gp}/R_p > 6$ – huge galaxy.
2^6	64	Negative or nan R_{gp}/R_p.
2^7	128	$(S/N)_w < 1$.
2^8	256	χ^2 per effective pixel >3.
2^9	512	Normed residuals <−20 somewhere.
2^{10}	1024	Normed residuals >20 somewhere.
2^{11}	2048	δu more than 10 arcsec from nominal.
2^{12}	4096	δv more than 10 arcsec from nominal.
2^{13}	8192	Failed to measure the FWHM of PSF or galaxy.
2^{14}	16 384	r-band SEXTRACTOR flag has 0×4 or above.
2^{30}	1073 741 824	No attempt at a fit was made due to cuts prior to running IM3SHAPE.

All images, including the co-add 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[25] to write chunks of images directly to disc 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`.[26] The code that generates the input object list, including cutout sizes, is part of the `meds` software library.

The MEDS data, including all of the images of each object observed in a single coadd tile, along with appropriate catalogue information, are stored in a single FITS file.

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[27] 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 of the meds file structure and the API for reading these files, we direct the reader to the `meds` repository URL.

APPENDIX B: DETAILS OF THE SHEAR CATALOGUES

B1 IM3SHAPE FLAGS

IM3SHAPE reports two kinds of flags. Table B1 lists 'error' flags, and Table B2 lists 'info' flags. For the most conservative treatment, users should select galaxies where both are zero. However, using `INFO_FLAG` > 0 may be appropriate in some cases.

[25] http://heasarc.gsfc.nasa.gov/fitsio/fitsio.html
[26] https://github.com/rmjarvis/deswl_shapelets
[27] https://github.com/esheldon/meds

Table B2. Info flags in the IM3SHAPE catalogue. Objects with non-zero INFO_FLAG may be acceptable depending on the scientific application.

Value	Decimal	Meaning
2^0	1	r-band SEXTRACTOR flagged with 0x1, indicating bright neighbours.
2^1	2	r-band SEXTRACTOR flagged with 0x2, indicating blending.
2^2	4	Mask fraction >0.5.
2^3	8	Model image <−0.01 somewhere.
2^4	16	$R_{gp}/R_p < 1.15$.
2^5	32	Radius >5 arcsec.
2^6	64	Centroid more than 0.6 arcsec from nominal.
2^7	128	χ^2 per effective pixel >1.25.
2^8	256	$R_{gp}/R_p > 3.5$ (very large galaxy).
2^9	512	Normed residuals <−2 somewhere.
2^{10}	1024	Normed residuals >2 somewhere.
2^{11}	2048	Declination below limit where we have good photometry.
2^{12}	4096	$(S/N)_w > 10^5$.
2^{13}	8192	Radius >10 arcsec.
2^{14}	16 384	$(S/N)_w < 10$.
2^{15}	32 768	Model image <−0.05 somewhere.
2^{16}	65 536	χ^2 per effective pixel <0.8.
2^{17}	131 072	More than 70 per cent of fitted flux is in masked region.
2^{18}	262 144	Model completely negative.
2^{19}	524 288	χ^2 per effective pixel >2.
2^{20}	1048 576	PSF FWHM >10 arcsec.
2^{21}	2097 152	Negative PSF FWHM.
2^{22}	4194 304	$R_{gp}/R_p < 1.1$.
2^{23}	8388 608	Centroid more than 1 arcsec from nominal.
2^{24}	16 777 216	Mask fraction >0.75.
2^{25}	33 554 432	One or more error flags is set.

B2 NGMIX FLAGS

The NGMIX catalogue 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 B3.

B3 File structure

There are three files available on the DESDM SVA1 Release web page[28] containing the final DES SV shear catalogues:

(i) `sva1_gold_r1.1_im3shape.fits.gz` is the IM3SHAPE catalogue. The relevant columns in this catalogue are listed in Table B4.

(ii) `sva1_gold_r1.0_ngmix.fits.gz` is the NGMIX catalogue. The relevant columns in this catalogue are listed in Table B5.

(iii) `sva1_gold_r1.0_wlinfo.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 catalogue, such as RA and declination, for convenience in using these catalogues without having to join them to the main DES object catalogue. Photometric redshift information is based on the SkyNet algorithm (Sánchez et al. 2014; Bonnett 2015; Bonnett et al. 2015). The columns in this catalogue are listed in Table B6.

[28] http://des.ncsa.illinois.edu/releases/sva1

Table B3. Error flag values in the NGMIX catalogue. Objects with non-zero ERROR_FLAG should be removed from any science analysis.

Value	Decimal	Meaning
2^0	1	There were no cutouts for this object.
2^1	2	PSF fitting failed for all epochs.
2^2	4	Not used.
2^3	8	Galaxy fitting failed.
2^4	16	Box size was larger than 2048.
2^5	32	Not used.
2^6	64	The $(S/N)_w$ of the PSF flux was lower than 4 in all bands.
2^7	128	Utter failure of the fitting. For this release, the flag was set when no valid guess for the fitters could be generated.
2^8	256	Not used.
2^9	512	No attempt was made for round measures because of prior failure.
2^{10}	1024	Round model could not be evaluated within the allowed region of parameter space.
2^{11}	2048	Fitting failed when trying to estimate S/N of T_r.
2^{30}	1073 741 824	No attempt of a fit was made due to other flags.

Table B4. The most relevant columns in the IM3SHAPE catalogue. Additional columns included in the catalogues are described in the documentation on the release web page.

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.
ERROR_FLAG	The error flag (cf. Table B1 in Appendix B1).
INFO_FLAG	The info flag (cf. Table B2 in Appendix B1).
SNR_W	The estimated $(S/N)_w$.
SNR_R	The estimated $(S/N)_r$.
MEAN_RGPP_RP	The mean value of R_{gp}/R_p among the different observations of the galaxy.

In addition to these three weak lensing files, the SVA1 release includes files for the Gold Catalogue, limiting magnitude maps, the good-region footprint, photo-z catalogues (Bonnett et al. 2015; including both point estimates and full PDFs), redMaPPer cluster catalogues (Rykoff et al. 2016), and redMaGiC galaxy catalogues (Rozo et al. 2015). See the release documentation page for more details about these other files.

Most users will want to select objects where `SVA1_FLAG == 0`. This selects the objects that we are confident are actually galaxies, and not either stars or some kind of spurious artefact in the data. See Table B7 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_FLAG` column is zero if

```
(ERROR_FLAG==0) & (INFO_FLAG==0) &
(SNR_W>15) & (MEAN_RGPP_RP>1.2)
```

The `NGMIX_FLAG` column is zero if

```
(ERROR_FLAG==0) & (0.4<ARATE<0.6) &
(SENS_1>0.0) & (SENS_2>0.0) &
(SNR_R>15) & (T_R/MEAN_PSF_T>0.15).
```

In each case, these select the galaxies which have been found to pass all of the null tests in Section 8. Users can thus

Table B5. The most relevant columns in the NGMIX catalogue. Additional columns included in the catalogues are described in the documentation on the release web page.

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 average of the two sensitivity estimates.
W	The recommended weight.
E_COV_*i_j*	($i, j \in \{1,2\}$) The covariance matrix of the shape estimate.
ERROR_FLAG	The error flag (cf. Table B3 in Appendix B2).
SNR_W	The estimated $(S/N)_w$.
SNR_R	The estimated $(S/N)_r$.
T	An estimate of the size, T, in arcsec^2.
T_R	An estimate of the size, T, that the object would have had if it were round.
MEAN_PSF_T	The mean measured size, T, of the PSF for the different exposures that went into the shear estimates for this galaxy.
ARATE	Acceptance rate of the MCMC chain.

Table B6. The columns in the info catalogue.

Column	Meaning
COADD_OBJECTS_ID	A unique id number of the object
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
PHOTOZ_BIN	The cosmological photometric redshift bin (0,1,2) as described in The Dark Energy Survey Collaboration et al. (2015)
MEAN_PHOTOZ	A point estimate of the photometric redshift from the SkyNet photo-z catalogue (Bonnett et al. 2015)
SVA1_FLAG	A flag indicating problematic galaxies (cf. Table B7)
IM3SHAPE_FLAG	A flag that is 0 if this object's shape in the IM3SHAPE catalogue is good to use; 1 if not.
NGMIX_FLAG	A flag that is 0 if this object's shape in the NGMIX catalogue is good to use; 1 if not.

Table B7. Values of the SVA1_FLAG in the info catalogue

Value	Decimal	Meaning
2^0	1	*i*-band SExtractor flag has bit 0×1 set, indicating bright neighbours.
2^1	2	*i*-band SExtractor flag has bit 0×2 set, indicating blending.
2^2	4	Modest Classification calls this object a star (bright_test or locus_test from Section 2.2).
2^3	8	Modest Classification calls this object junk (faint_psf_test from Section 2.2).
2^4	16	In region with high density of objects with 'crazy colours' (i.e. any of the following: $g-r<-1$, $g-r>4$, $i-z<-1$, or $i-z>4$).
2^5	32	In region with a high density of objects with large centroid shifts between bandpasses.
2^6	64	Near a 2MASS star with $J_M<12$. The mask radius is flux dependent, up to 120 arcmin for the brightest stars.
2^7	128	Large offset in *g*- and *i*-band windowed positions.
2^8	256	Object was not measured by NGMIX.
2^9	512	Likely star according to NGMIX $T+\sigma_T<0.02$ square arcsec.
2^{10}	1024	Very low surface brightness according to NGMIX measurements.
2^{11}	2048	Object does not satisfy good measurement flags in NGMIX.
2^{12}	4096	Object does not have a valid magnitude in all *g*,*r*,*i*,*z* bands. (That is, at least one of them is invalid.)

select galaxies with good shear estimates by simply selecting IM3SHAPE_FLAG==0 or NGMIX_FLAG==0 as appropriate.

[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, PO 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, D-85748 Garching, Germany*
[11]*Faculty of Physics, Ludwig-Maximilians University, Scheinerstr. 1, D-81679 Munich, Germany*
[12]*Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[13]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, USA*
[14]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[15]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, D-81679 München, 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]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa*
[23]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[24]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[25]*Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK*
[26]*Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[27]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[28]*Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA*
[29]*National Center for Supercomputing Applications, 1205 West Clark St, Urbana, IL 61801, USA*
[30]*Institut de Ciències de l'Espai, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, E-08193 Bellaterra, Barcelona, Spain*
[31]*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*
[32]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[33]*Australian Astronomical Observatory, North Ryde, NSW 2113, Australia*
[34]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, CEP 05314-970, São Paulo, SP, Brazil*
[35]*Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*
[36]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[37]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton BN1 9QH, UK*
[38]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
[39]*Department of Physics, University of Illinois, 1110 W. Green St, Urbana, IL 61801, USA*

This paper has been typeset from a TEX/LATEX file prepared by the author.