Cosmic shear measurements with Dark Energy Survey science verification data

Article  (Accepted Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/65679/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk
Cosmic Shear Measurements with DES Science Verification Data

(The DES Collaboration)

1 Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA
2 Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, 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 and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
5 Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
6 Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany
7 Universität-Gesamthochschule Paderborn, Fakultät für Physik, 33095 Paderborn, Germany
8 Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland
9 Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
10 Institut de Física d’Altes Energies, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain
11 Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
12 Department of Physics, University of Chicago, Ann Arbor, MI 48109, USA
13 Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
14 Institut de Ciències de l’Espai, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, 08193 Bellaterra, Barcelona, Spain
15 Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
16 Kavli Institute for Particle Astrophysics, University of California, Menlo Park, CA 94025, USA
17 Korea Astronomy and Space Science Institute, Yusong-gu, Daejeon, 305-348, Korea
18 Department of Physics, The Ohio State University, Columbus, OH 43210, USA
19 Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
20 Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile
21 Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa
22 Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA
23 Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
24 Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
25 CNRS, UMR 7095, Institut d’Astrophysique de Paris, F-75014, Paris, France
26 Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d’Astrophysique de Paris, F-75014, Paris, France
27 Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
28 Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
29 National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA
30 George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
31 Department of Physics, Ludwig-Maximilians University, Scheinerstr. 1, 81679 Munich, Germany
32 Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
We present measurements of weak gravitational lensing cosmic shear two-point statistics using Dark Energy Survey Science Verification data. We demonstrate that our results are robust to the choice of shear measurement pipeline, either ngmix or IM3SHAPE, and robust to the choice of two-point statistic, including both real and Fourier-space statistics. Our results pass a suite of null tests including tests for B-mode contamination and direct tests for any dependence of the two-point functions on a set of 16 observing conditions and galaxy properties, such as seeing, airmass, galaxy color, galaxy magnitude, etc. We furthermore use a large suite of simulations to compute the covariance matrix of the cosmic shear measurements and assign statistical significance to our null tests. We find that our covariance matrix is consistent with the halo model prediction, indicating that it has the appropriate level of halo sample variance. We compare the same jackknife procedure applied to the data and the simulations in order to search for additional sources of noise not captured by the simulations. We find no statistically significant extra sources of noise in the data. The overall detection significance with tomography for our highest source density catalog is 9.7σ. Cosmological constraints from the measurements in this work are presented in a companion paper [1].

I. INTRODUCTION

Cosmic shear, the weak gravitational lensing of galaxies due to large-scale structure, is one of the most statistically powerful probes of Dark Energy, massive neutrinos, and potential modifications to General Relativity [2, 3]. Due to its powerful potential as a cosmological probe, many ongoing and future surveys (Kilo-Degree Survey: KiDS, Hyper Suprime-Cam survey: HSC, the Dark Energy Survey: DES, the Large Synoptic Survey Telescope: LSST, Euclid, and WFIRST) will employ cosmic shear as one of their principle cosmological probes. Cosmic shear two-point measurements, in their simplest form, are made by correlating the shapes of many millions of galaxies as a function of their separation in angle. Additionally, if the galaxies can be separated as a function of redshift, then tomographic cosmic shear measurements can be made by cross-correlating galaxies at different redshifts, which can probe the evolution of large-scale structure. The galaxies themselves have intrinsic shapes that are an order of magnitude larger than the cosmic shear signal, which means that cosmic shear measurements involve extracting small correlations from a large, shape noise-dominated background. Competitive cosmological constraints from cosmic shear will require of order percent level or better measurements at all steps of the analysis, from shear measurement to the measurements of cosmic shear two-point functions (see, e.g., Weinberg et al. [4] or Kilbinger [5] for a review).

Cosmic shear was first detected in 2000 [6, 7]. The most recent results have detected correlated shapes on scales from a few to 60 arcminutes from the Deep Lens Survey [10], the Sloan Digital Sky Survey [11, 12], KiDS [13] and the Canada-France-Hawaii Legacy Survey [14], including in 6 redshift bins [15]. Future cosmic shear measurements will be very high signal-to-noise and over much larger survey areas, yielding a wealth of cosmological information.

Cosmic shear measurements are challenging for a variety of reasons. First and foremost, shear measurements are subject to biases that can arise from a number of sources. These biases are usually split into additive and multiplicative components. Sources of additive biases include inaccuracies in the modeling of the point spread function (PSF), inaccuracies in correcting for the effect of the PSF on galaxy images, astrometric errors, and contaminating flux from nearby galaxies. Multiplicative biases can arise from the effects of noise on the shear measurement process, incorrect estimates of the size of the PSF, and, for model-fitting methods, mismatches between an object’s true underlying structure and the model employed in the shear measurement process. Additionally, many modern shear measurement methods require accurate estimates of the distribution of galaxy
shapes and profiles in the absence of lensing to either serve as priors in the extraction of shapes from the data or to directly make corrections to the data. These priors can be estimated from high-resolution Hubble Space Telescope imaging, but must be matched to the observational sample under consideration.

Significant computational and scientific challenges in cosmic shear measurements remain, even in the presence of perfect shear measurements. The cosmic shear field is the result of lensing by the non-linearly evolved matter density field. Accurate predictions for the non-linear matter power spectrum, even just for pure dark matter models, are computationally expensive and are needed at every point in parameter space in order to extract cosmological parameters. Emulators, like the Coyote Universe [16], have solved this problem for typical cosmologies and Dark Energy models, but neglect important physical effects, like galaxy formation, on the matter power spectrum. Additionally, some physical effects of galaxy formation break the assumption that galaxies are randomly oriented in the absence of lensing. These effects, called intrinsic alignments, can introduce correlations in the shapes of galaxies that are not due to lensing, complicating the interpretation of cosmic shear measurements [see, e.g., 17, 18]. Furthermore, even if the mean signal can be modeled properly, the covariance matrix of cosmic shear measurements is dominated by sample variance, requiring either extensive suites of numerical simulations or complicated halo model calculations. The (mis-)estimation of photometric redshifts (photometric simulations or complicated halo model calculations. Appendix A describes alternate two-point estimators besides the real-space correlation functions used for the bulk of this work. We discuss the estimation and validation of our covariance estimation in Section V. Then, we describe our suite of null and consistency tests of our measurements in Section VI. Finally, we conclude in Section VII. The shear correlation functions and simulation covariance matrices from this work are available as online supplementary material with this paper.

II. DATA

The DES SV data with weak lensing measurements consists of 139 square degrees of five-band imaging with roughly 7 exposures per band on average [21, 22]. The depth of the data is somewhat shallower than the expected ∼10-exposure average depth of the DES five-year data. The basic reductions and co-add source detection were done with the DES data management (DESDM) system as described in Desai et al. [25] and Gruendl et al. (in preparation). We use the shear measurements from Jarvis et al. [19] performed on the DES SV Gold sample of galaxies (Rykoff et al., in preparation). For more information on the shear measurements and recommended cuts, we refer the reader to Jarvis et al. [19].

The shear measurement pipelines and photo-28 used in this work are described below for completeness. We use the “reduced shear” ellipticity definition [24]. Finally, note that the two shear measurement codes used in this work are not identical, employing different cuts and different parts of the DES SV data. Thus they have different overall source number densities and photometric redshift distributions. These differences, which we expect to be smaller in future DES analyses (see Jarvis et al. [19]), have no effect on the major conclusions of this work and are in fact important in verifying the robustness of our results.

A. Shear Measurement Pipeline 1: ngmix

The NGMIX pipeline [27] uses sums of Gaussians to represent simple galaxy models [28]. The model parameters of each object are sampled using Markov Chain Monte Carlo (MCMC) techniques applied to a full likelihood which forward models the galaxy and its convolution with the PSF. The total likelihood for each object
is a product of the likelihoods of the individual images of each object. The $r$-, $i$- and $z$-bands are fit simultaneously with the same model shape, but different amplitudes. The samples of the likelihood are then used with the \texttt{lenstool} algorithm \cite{2007ApJ...666..567K} to measure the shear of each object using a prior on the intrinsic distribution of shapes from the GREAT3 \cite{2012MNRAS.421..537C} release of the COSMOS galaxy sample. The final effective source number density of the \texttt{NGMix} catalog is $\approx 6.1$ galaxies per square arcminute\footnote{We use the following definitions of effective source density $n_{\text{eff}}$ and the effective shape noise per component $\sigma_{\text{SN}}$, which are appropriate for the two-point function estimators employed in this work. $n_{\text{eff}} = \left( \sum w_i n_i \right) / \left( \sum w_i^2 s_i^2 \right)$ and $\sigma_{\text{SN}}^2 = \left( \sum w_i^2 (e_{i1}^2 + e_{i2}^2) \right) / \left( 2 \sum w_i^2 s_i^2 \right)$ where $w_i$ are the weights, $s_i$ are the sensitivities, $e_i$ are the shear components, $\Omega$ is the survey areas and the index $i$ runs over all of the galaxies. \url{https://bitbucket.org/joezuntz/im3shape}}. Each source has an associated weight and we use the average sensitivity over both directions, as described in Jarvis \textit{et al.} \cite{2014ApJ...796...29J}. We use a set of 126 mock catalogs to compute the covariance matrix of the shear correlation functions, E/B-mode statistics, power spectra and null statistics described in the following sections. These mock catalogs are constructed from seven sets of simulations consisting of three N-body light cones pieced together along the line of sight. We use 1050 $h^{-1}\text{Mpc}$, 2600 $h^{-1}\text{Mpc}$ and 4000 $h^{-1}\text{Mpc}$ boxes with 1400$^3$, 2048$^3$ and 2048$^3$ particles respectively. We use a flat, $\Lambda$CDM model with $\Omega_m = 0.286$, $\Omega_k = 0.714$, $n_s = 0.96$, $h = 0.7$, $\Omega_b = 0.047$, $w = -1$ and $\sigma_8 = 0.820$. The initial conditions are generated at redshift 49 with 2LPTic, a second-order Lagrangian perturbation theory initial conditions generator \cite{2001ApJ...556..740K} using linear power spectra from the \texttt{CAMB} Boltzmann code \cite{2002PhRvD..66j3511L}. The N-body evolution is computed with an efficient dark-matter-only version of the \texttt{Gadget-2} code \cite{2005MNRAS.364.1105S}. We have implemented our own on-the-fly light cone generator directly into the \texttt{Lahab} code (Busha et al. in preparation). We produce a full-sky light cone which formally replicates the N-body box eight times. However, each final simulation covers only one octant of the full-sky, $\approx 5,000$ square degrees, eliminating the replications. As the DES SV area with weak lensing measurements is only 139 square degrees, we divide each simulation into 18 different pieces using the observed SV mask to construct 126 total mock catalogs. This procedure has the advantage of properly computing the halo sample variance contributions to the lensing covariance matrices due to the fact that each patch is embedded in the large-scale modes of the box.

We place lensing sources randomly in angle with in the DES SV mask (see Jarvis \textit{et al.} \cite{2014ApJ...796...29J} for the details of the mask), and with the redshift distribution of the tomographic bins defined above. Then the weak lensing shear for each source is computed using the \texttt{CALCLENs} ray-tracing code \cite{2009ApJ...700.1739K}. In this application of \texttt{CALCLENs}, we use the pure spherical harmonic transform version with \texttt{Nside}=8192. Appendix \cite{2014ApJ...796...29J} presents tests of the underlying simulations in comparison to simple expectations from fitting functions to the matter power spectrum. We find that the simple expectations from matter power spectrum fitting functions agree with the simulation to within sample variance, but that some resolution...
FIG. 1. The measured shear correlation functions $\xi_{+/-}$ for a single tomographic bin for the NGMIX shape catalog (left) and IM3SHAPE shape catalog (right). The single tomographic bin corresponds to redshift distribution shown in Figure 3, $z \approx 0.3-1.3$. Note that the redshift distributions of the two catalogs are not identical, so that the shear correlation functions are not expected to match. A detailed comparison of the two catalogs is described in Section VI B. Negative measurements are shown as upper limits. The error bars show the $1\sigma$ uncertainties from the mock catalogs with the appropriate level of shape noise for each shear pipeline. The black solid lines show the predictions from a flat, $\Lambda$CDM model described in Section III — not chosen to fit the data.

issues remain on small scales. Note, however, that these small scales are excluded from the companion cosmological analysis [1] and that despite the resolution issues, we find excellent agreement between the covariances computed from the mock catalogs and the halo model, as discussed below. Thus for purposes of computing covariance matrices, the mock catalogs we have constructed are sufficient. Future work may require higher-resolution shear fields for covariance estimation.

Finally, we generate the shape noise and other properties in the mock by randomly drawing from the observations separately for each tomographic bin. Importantly, we draw the intrinsic shape of each mock shear source separately from its other properties, like signal-to-noise, size, etc. Properties which have intrinsic spatial dependence in the survey (e.g. seeing, airmass, etc.) are drawn from the nearest real galaxy to each mock galaxy. See Section IV C for more details. These procedures randomise the shear field in the data and ensure that the mock catalogs have no correlations between the systematic parameters and the shear field.

IV. MEASUREMENTS OF COSMIC SHEAR TWO-POINT STATISTICS

In this work, we focus on cosmic shear measurements made with two-point statistics, which are detailed in the following sections. A companion paper [1] presents the associated cosmological parameter constraints using these measurements, which use the real-space two-point correlation functions as the fiducial two-point estimator. We summarize results from alternate estimators in Section IV C and Appendix A. Note that although the choice of which two-point statistic to use is somewhat arbitrary, the companion cosmological analysis of this data [1] demonstrates that the exact choice of two-point statistic does not change the cosmological parameter constraints from this data in a statistically significant way.

A. Real-space Two-point Function Estimators

We follow Miller et al. [11] and estimate the two-point functions with

$$\xi_{\pm} = X_{\pm} \pm X_{\times}$$

$$X_{\pm/\times} = \frac{\sum_{i,j} w_i w_j (e - c)_{i+,\times} (e - c)_{j+,\times}}{\sum_{i,j} w_i w_j s_i s_j}$$

(1)
where $i, j$ index the galaxies in the two sets we are correlating. Here $e_{+}/\times$ are the estimated shears from the lensing analysis projected into the + (tangential) and $\times$ (cross) components rotated into the reference frame connecting each pair of galaxies $\{i, j\}$ in the sum. The $w_i$ are weights applied to each galaxy (typically inverse variance weighting; see Sec. II for each lensing code). The $s_i$ are multiplicative noise bias and/or lensing sensitivity corrections that are applied to the shears. We follow Miller et al. [41] and apply these corrections to the entire population of shears as opposed to applying them to each shear individually. We compared several different methods for incorporating the sensitivities into the two-point function estimator and find that they differ by at most $\sim 2\%$. The $c_i$ are the additive bias corrections used for Im3Shape and are identically zero for ngmix per the definition of the lensfit method [29]. Finally, we use TreeCorr\textsuperscript{10} [42] to compute the shear correlation functions.

B. Real-space Correlation Functions

The real-space correlation functions without tomography are shown in Figure 1. We show ngmix on the left and Im3Shape on the right, with $\xi_+$ in the top rows and $\xi_-$ in the bottom rows. Negative measurements are shown as upper limits. The redshift distribution of sources for the non-tomographic analysis is shown in the top panel of Figure 3 for the SkyNet code. It appears in Figure 1 that the $\xi_+$ correlation function may approach a constant value at large scales. Interestingly, Jarvis et al. [19] find that the mean shear across the survey for ngmix and Im3Shape is $\approx 7 - 10 \times 10^{-4}$. This level of mean shear would produce a constant floor in the shear correlation functions of $\approx 5 - 10 \times 10^{-7}$. For the DES SV survey, the root-mean-square mean shear just due to shape noise and cosmic variance is $\approx 4 \times 10^{-4}$. Thus it is not clear if this feature is an indication of systematic effects or a few sigma fluctuation in the mean shear due to a real physical effect. However, in the cosmological analysis of this data, all $\xi_+$ data points above 60 arcminutes were cut to avoid systematics in the PSF models [1, 19]. Thus we do not explore this issue further in this work.

We generate estimates of the 1σ uncertainties for each measurement by computing the covariance of the two-point functions over the simulation mock catalogs described in Section III. These mock catalogs are built sep-

\textsuperscript{10} https://github.com/rmjarvis/TreeCorr
FIG. 3. The estimated redshift distributions from SkyNet for the NGMIX shape catalog (left) and the IM3SHAPE shape catalog (right). The full \( n(z) \) for objects with mean redshifts in the redshift range \( 0.3 < \bar{z} < 1.3 \) (top) and the \( n(z) \) for three tomographic bins (bottom) are shown. The redshift distributions are estimated by summing and rescaling the photometric redshift probability distributions for each galaxy in the tomographic bin using the weights applied to the shear catalog separately for each shear catalog in order to match the non-tomographic redshift distribution of the sources. The correction factor described in Hartlap et al. [43] is then applied to produce an unbiased estimate (see Section V A for a further discussion of the statistical properties of the covariance matrix estimate from the mock catalogs). The significance of the resulting measurement is then calculated from this covariance as

\[
S/N = \frac{\xi_{\text{data}} C^{-1} \xi_{\text{model}}}{\sqrt{\xi_{\text{model}} C^{-1} \xi_{\text{model}}}},
\]

(2)

where \( C^{-1} \) is the inverse covariance matrix estimated from the mock catalogs, \( \xi_{\text{data}} \) is the vector of real-space shear two-point function measurements from the data, and \( \xi_{\text{model}} \) is the vector of real-space shear two-point function measurements predicted from the cosmological model given above in Section III. This quantity corresponds to the signal-to-noise of a least-squares estimate of a scaling parameter comparing our measurements to the theoretical model. This signal-to-noise measure will be an underestimate if the model employed is not well matched to the data. However, given the good match of our fiducial model to the data as shown in Figures 1 and 2, the degree to which the signal-to-noise is underestimated is small in this case. We use the COSMOSIS package [11] by Zuntz et al. [44] to compute the shear correlation functions with the Takahashi et al. [45] non-linear power spectrum fitting function. See the companion paper [1] presenting cosmological constraints from these measurements for additional details on the model correlation function \( \xi_{\text{model}} \) computation. The covariance matrix has been validated through comparisons to both a detailed halo model prediction and jackknife estimates in single mock patches versus the survey data, which are discussed in detail in the Section V. We find non-tomographic cosmic shear detections at 6.5\( \sigma \) and 4.7\( \sigma \) significance for NGMIX and IM3SHAPE respectively.

Figure 2 shows the full three-bin tomographic shear correlation function measurements for NGMIX on the left and IM3SHAPE on the right. The redshift distributions of the three tomographic bins for the SkyNet code are given in the lower panels of Figure 3. In order to compute the covariance matrix of these measurements, we use the same procedure in the mock catalogs as for the non-tomographic case, except that we use the tomographic redshift distributions to assign the mock galaxies to different tomographic bins. We additionally draw the shape noise in the mock from only the galaxies in the data in the same tomographic bin. We find overall tomographic cosmic shear detections of 9.7\( \sigma \) and 7.0\( \sigma \) for NGMIX and IM3SHAPE, respectively. Note that the NGMIX catalog has more sources and extends to slightly higher redshift on average, yielding higher significance detections of cosmic shear. We have chosen three tomographic bins as a compromise between gaining signal-to-noise in the data and having too many data points in order to use the mocks to compute the covariance matrix of the data.

In Figures 1 and 2, the solid black line shows the expected amplitude and shape of the shear correlation functions in the cosmological model given above. This curve is not a fit, and is merely presented as a reference for comparison. Due to the fact that the two catalogs have different redshift distributions, a direct comparison of the shear correlation functions between the two catalogs is not possible without further work matching the two cat-

\[11\] https://bitbucket.org/joezuntz/cosmosis
alogs and accounting for the shared shape noise, sample variance, and image noise between the two catalogs. This matched comparison is described further in Sec. VTB.

C. Alternative Two-point Statistics

In Appendix A we describe results from two alternative two-point statistics of the shear field. These include the methods of: (i) Becker and Rozo [46], which use a weighting of the real-space correlation estimates to construct efficient estimates of the $C_T$ values and (ii) a second estimation of the spherical harmonic shear power spectrum using PolSpice [47, 48]. Note that these estimators weight the data at different angular scales differently than the default two-point correlation functions so that we do not expect to get identical results in terms of the significance of the cosmic shear detection. We do find detections of cosmic shear that are consistent with the conventional real-space estimators we use by default, indicating no strong preference for any given estimator. Tests of B-mode statistics from these estimators are discussed in Sec. VTA where we again find consistency between different two-point function estimation methods.

V. ESTIMATING AND VALIDATING THE COVARIANCE MATRIX

In this section, we present our covariance matrix and a set of validation tests. The fiducial covariance matrix for our measurements is estimated from the mock catalogs presented in Section III. First we compare the covariance matrix from the mock catalogs to halo model computations. Second, we compare jackknife covariances in the data to the jackknife covariance computed from the mock catalogs. This procedure allows us to look for additional sources of noise and correlations in the data that are not present in the mock catalogs.

A. Simulation and Halo Model Comparison

We compare the covariance matrix computed from the simulations to that obtained from a halo model in Figure 4. The simulation-based covariance matrix is computed by populating the mock catalogs with shear sources as described above in Section III and then computing the covariance of the measurements performed on the full ensemble of mock catalogs. The halo model covariance was computed with the CosmoLike covariance module (see Eifler et al. [49] and Krause et al. [50] for details). Further details of our halo model computation and the full tomographic covariance matrix are given in Appendix C. Briefly, we include the Gaussian, non-Gaussian and halo sample variance terms [e.g., 51] and compute the halo model covariance at the same cosmology and with the same redshift distribution as was used in the mock catalogs.

We compare the general structure of the mock (upper triangle) and halo model (lower triangle) covariance in the left panel of Figure 4, which shows part of the correlation matrix. Here we have shown a subset of the full set of tomographic bin combinations. The full correlation matrix is shown in Appendix C. The right panel compares the amplitude of the two covariances by plotting the variance. Overall, we find good agreement in both structure and amplitude.

We quantitatively test the agreement using a Fisher matrix computation. We compute the expected error on the degenerate parameter combination $\sigma_s(\Omega_m/0.3)^{0.5}$, where $\sigma_s$ is the RMS amplitude of the linear matter power spectrum at redshift zero in a top hat window of $8 \, h^{-1}\text{Mpc}$ and $\Omega_m$ is the matter density in units of the critical density at $z = 0$. This combination of parameters is typically the best constrained by low-redshift cosmic shear data sets like the DES SV data. The exact degeneracy is computed in the companion cosmological constraints paper to this work [1]. We use the standard Fisher formalism for cosmic shear [see, e.g., 52] and the same cosmological model as described above. We vary only the spectral index $n_s$, $\sigma_s$ and $\Omega_m$ in the Fisher matrix.

We find that the error bars on $\sigma_s(\Omega_m/0.3)^{0.5}$ from the halo model and mock covariances agree to approximately 10% without tomography, with the halo model yielding larger parameter uncertainties. When repeating the same exercise with tomography, we find a larger, $\approx 35\%$ disagreement in the error bars, with the mocks yielding larger errors. However, we expect fluctuations in the uncertainties in parameters computed with the simulations due to the finite number of realizations used for the covariance computation. Dodelson and Schneider [53] estimate that this effect, in the Gaussian limit, increases the variance in the parameter estimates by a factor of

$$\alpha = 1 + \frac{(N_d - N_p)(N_s - N_d - 2)}{(N_s - N_b - 1)(N_s - N_b - 4)}$$

where $N_d$ is the number of data points, $N_s$ is the number of simulations and $N_p$ is the number of parameters. This factor is $\approx 1 + N_d/N_s$ in the limit that $N_s \gg N_d \gg N_p$. Thus we expect a fractional uncertainty in the parameter uncertainties of $\approx \sqrt{\alpha} - 1$. In our case with tomography, $N_d = 72$, $N_s = 126$ and $N_p = 1$. With these numbers, we get that the fractional uncertainty in the parameter uncertainty is $\approx 118\%$. Thus the disagreement of $\approx 35\%$ we find with the halo model with tomography is not statistically significant. Without tomography, we find a fractional uncertainty in the uncertainty of $\approx 56\%$, again indicating consistency.

Importantly, these numbers are the fractional uncertainty in the uncertainty. For parameter estimates, the
fractional increase in the uncertainty on the parameter, equal to $\sqrt{\alpha}$, is the relevant quantity. For tomography, this fractional increase is $\approx 55\%$ and without tomography it is $\approx 15\%$. Furthermore, we have assumed that the tomographic analysis uses all 72 data points. As described in The Dark Energy Survey Collaboration et al. [1], only 36 of the 72 data points are used for tomography, bringing the fractional increase in the error due to the finite number of realizations down to only $\approx 18\%$. Similar cuts are made for the non-tomographic analysis, using only 16 of the 30 data points. This number of data points results in a fractional increase of the parameter uncertainties of only $\approx 7\%$ for the non-tomographic analysis.

B. Jackknife Comparisons to Data

While our mock catalogs include both sample variance and shape noise contributions, any spatially varying systematic effects, like errors in the shear calibration, should
be included in the covariance matrix of the shear correlation functions as well. To search for these potential effects, we use the jackknife covariance matrix of the shear correlation functions as a statistic to be compared between the data and the mock catalogs. Any additional sources of noise in the data, which are captured by the spatial scale of our jackknife regions, will show up as a difference between the jackknife covariance as computed in the data versus the mock catalogs.

We estimate the jackknife covariances from the data and our mock catalogs as follows. We divide both the mock catalog and data into 100 spatial sub-regions, employing the k-means algorithm.13 These regions are then used to perform jackknife resampling. For the details of jackknife covariance estimation for cosmic shear correlation functions, we refer the reader to a (technical) companion paper where these choices are examined in further detail (Friedrich et al. [54], see also Norberg et al. [55] for an application to galaxy clustering). We use the standard jackknife scheme, where all of the shear sources in an entire subregion are removed for each jackknife resampling, which is called the galaxy-jackknife in Friedrich et al. [54].

Note that we are not comparing jackknife covariances with the true covariances, but rather simply the covariance in the shear correlation function across the DES SV survey to the same statistic computed with the mock catalogs. Thus the absolute correctness of the jackknife covariance matrix is not an issue for our test, since it is just a statistic that is sensitive to the effects for which we wish to search. The performance of empirical covariance measures for cosmic shear surveys is explored in Friedrich et al. [54].

The comparison of our jackknife procedure between the mocks and the data is shown in Figure 5. Here we plot the correlation matrix of the averaged jackknife covariance from the 126 mock NGMIX catalogs (left panel, on the bottom right) and the same computation in the DES SV data (on the top left). The right panel compares the diagonal elements of the jackknife covariance for \(\xi_+\) and \(\xi_-\) when averaged over 126 mock catalogs and when computed from the data for NGMIX. Using the Fisher matrix procedure described above, we find that the error on \(\sigma_8(\Omega_m/0.3)^{0.5}\) from the data jackknife covariance matrix agrees with the mean of the ensemble of errors on this parameter from the mock jackknife covariances to within one standard deviation of the error over the ensemble. Thus we conclude that there are no statistically significant sources of additional variance in the data compared to the mock catalogs.

VI. TESTS FOR RESIDUAL SYSTEMATIC ERRORS IN THE COSMIC SHEAR SIGNAL

Systematic errors in shear measurements can be from a wide array of sources ranging from telescope optics and observing conditions to details in the modeling, measurement, and calibration of shapes. The development of tests to identify potential systematic errors is critical to verifying accurate measurement of cosmic shear. Toward this end, we devise a set of tests that should produce a null result when applied to true gravitational shear. The measurement of a significant non-zero result is then an indication of unresolved systematic errors in the shear catalog that could bias measurements.

The DES SV shear catalogs have passed a rigorous set of both traditional and novel null tests that lay the groundwork for validating the precise measurements that will be made with ongoing DES measurements during the main survey observing period. These tests are performed both at the catalog level and during the process of validating specific measurements based on the shear catalogs. We describe the methodology and results of both traditional and new null tests for sources of potential systematic errors in both the non-tomographic and tomographic measured cosmic shear signal in the next two sections.

Catalog-level tests were performed by Jarvis et al. [19] cf. their Section 8] and included tests for additive systematic errors related to spatial position, the PSF, and galaxy properties. These tests included the cross-correlation of the galaxies and the PSF. No significant additive systematic errors were found, and they put upper limits on the potential additive systematic contribution to \(\xi_+\) in their Section 8.7. In addition, the overall multiplicative bias of the shear estimates was tested with simulations in Jarvis et al. [19]. Jarvis et al. [19] concludes that both catalogs are consistent with having small overall multiplicative bias, but due to uncertainties in their ability to constrain this value, they suggest marginalizing over a prior on the multiplicative bias with a standard deviation of 0.05 (see Equation 8-12 of Jarvis et al. [19]). This multiplicative systematic is treated in the cosmological analysis of this data [1], where it contributes to an increase in the uncertainties on the final cosmological parameters constrained with this data.

A. B-mode Measurements

The cosmic shear field can be characterized by E- and B-modes which differ in parity. At first-order in the gravitational potential in General Relativity, cosmic shear produces a pure E-mode field [see, e.g. 56]. However, contaminating signals, like that from the telescope point-spread function, tend to contain both E- and B-modes. Thus one of the first suggested tests of cosmic shear detections was verifying that the B-mode signal is consis-

---

FIG. 6. Tomographic B-modes in DES SV data for NGMIX (left) and IM3SHAPE (right). The error bars are calculated from the simulation realizations using the shape noise appropriate for each catalog. The tomographic bins correspond to those shown in Fig. 3 and are labeled from 1 to 3, increasing with redshift. Thus, panel ‘3-2’ shows the cross-correlation between the highest and middle redshift bins. The total $\chi^2$/d.o.f., accounting for the correlations between the points in each panel, for NGMIX is 62.5/60 and for IM3SHAPE is 41.2/60.

Small levels of B-modes are produced at second order in the gravitational potential, but these are small enough not to spoil the null test [see, e.g., 58, 59].

Many methods have been suggested for B-mode estimation [e.g., 60–65]. Here we use the estimators from Becker and Rozo [46], which estimate bandpowers using linear combinations of the shear two-point functions that optimally separate E- and B-modes [40]. These estimators are

$$E = \frac{1}{2} \left[ \sum f_+ \xi_+ + \sum f_- \xi_- \right]$$

$$(3)$$

$$B = \frac{1}{2} \left[ \sum f_+ \xi_+ - \sum f_- \xi_- \right]$$

$$(4)$$

where the sums run over the angular bins of the shear two-point functions. The weight vectors $f_+/-$ are chosen to simultaneously minimize E- to B-mode mixing while also producing compact band-power estimates in Fourier-space. See Appendix A for more details.

In Figure 6, we show a measurement of the tomographic B-mode signal using the Becker and Rozo [46] band-powers. We find no statistically significant B-mode contamination, with a total $\chi^2$/d.o.f. for NGMIX of 62.5/60 and for IM3SHAPE of 41.2/60. The error bars in this case are computed using the mock catalogs above. In Appendix A, we verify this conclusion by computing a complementary measurement of the non-tomographic B-mode signal using an alternate estimation of the spherical harmonic shear power spectrum. We find the B-modes from this alternate technique are consistent with zero with $\chi^2$/d.o.f. = 4.5/7 for NGMIX and 6.3/7 for IM3SHAPE. Finally, note that Becker and Rozo [46] band-power measurements of the non-tomographic B-mode signal are presented in Jarvis et al. [19] using the methods and mock catalogs of this work. The non-tomographic B-mode measurements were again found to be consistent with zero, with $\chi^2$/d.o.f. = 22.3/20 for NGMIX and 16.1/20 for IM3SHAPE.

B. Consistency Between the Shear Pipelines

We further test for consistency between the shear catalogs split into tomographic bins by selecting only sources which pass the selection cuts for both codes. For this subset of sources, we then compare the shear auto- and cross-correlation functions for each bin. Due to the fact that the two catalogs have the same sample variance, have similar shape noise and have correlated shear measurement errors, the error bars on the difference between the two correlation functions is much smaller than that on the correlation functions themselves. We account for this effect by constructing mock catalogs where a given mock galaxy is assigned its shape noise for each shear...
FIG. 7. Difference over error in the tomographic correlation functions for matched shear catalogues from \textsc{ngmix} and \textsc{im3shape}. We show \textsc{im3shape} minus \textsc{ngmix}. The total $\chi^2$/d.o.f. accounting for all correlations is 46.8/72.

This comparison is shown in Figure 7 for the shear correlation function for \textsc{im3shape} minus \textsc{ngmix}. We find that the shear correlation functions from the codes are statistically consistent over the full range of scales from 2 to 300 arcminutes, giving a $\chi^2$/d.o.f. $= 46.8/72$. Finally, note that this test is similar to the differenced shear correlation function test presented in Section 8.6 of Jarvis et al. [19]. For their test, they examine the shear correlation function of the difference in the \textsc{ngmix} and \textsc{im3shape} shear estimates using the matched catalogs. They find that below $\approx$3 arcminutes, the catalogs do not meet the requirements for additive systematic errors, set by the expected precision of the cosmological constraints. The test presented in this work is generally less sensitive, but complementary, to the differenced shear correlation function.

C. Two-Point Null Tests

Even with a carefully chosen set of null tests at the catalog level, it is still possible that systematic errors, which can be due to complex interplays between different aspects of data and analysis, may influence the cosmic shear measurement. To test for any uncorrected systematic errors remaining in the measured cosmic shear signal, we attempt to measure the variation in $\xi_+$ as a function of survey and galaxy properties that may be correlated with sources systematic errors. For each survey or galaxy property, the shear data is split in half, and the correlation functions of each half are compared. We use a reweighting method to ensure that the redshift distribution of each half is the same in order to remove any cosmological dependence from this null test. If the photo-zs and shear measurements are correct, then the shear correlation functions of the two halves should be consistent to within the noise of the shear measurements and the redshift reweighting. If they are not, this would indicate either uncorrected systematics, selection effects from the split, or non-shear differences in the two halves such as intrinsic alignments.

Due to the fact that each half is drawn from the same area in many cases, the standard error bars computed for the shear correlation functions are not correct for this test. We instead use the mock catalogs described above to compute the error on the difference between the two halves relative to the full sample, accounting for shared sample variance, as described below. It is important to note that this is a simultaneous test of both the photometric redshifts and the shear calibrations. This feature is in fact desirable because both of these quantities can contribute to biases in the shear correlation functions. We have used both the survey property maps described by Leistedt et al. [66] and also properties directly produced by the shape measurement codes. The 16 various systematic parameters are described in Table 1. Finally, Jarvis et al. [19] found that making cuts on signal-to-noise and size could lead to a selection bias in the population of shear values due to preferentially selecting galaxies that look more or less like the PSF. We attempt to minimize this problem by using the “round” measures of signal-to-noise, size and surface brightness.

1. Methodology

The galaxies in each half-sample must be reweighted so that the total $n(z)$, computed from summing the individual $p(z)$ for each galaxy according to its weight, matches between the two half-samples. Matching the redshift distributions of the two halves removes any cosmological dependence in each null test. For the data, the extra weights are computed using Ridge Regression (or Tikhonov regularization) [67]. We use the Ridge Regression algorithm to solve for an additional weight for each galaxy, which when used with the shear measurement weights described in Section II to compute the $n(z)$, produces a matching redshift distribution between the two half-samples. The Ridge Regression algorithm solves the linear least-squares problem with an additional regular-

---

15 We have completed this test for the ratio of the shear correlation functions and without tomography, finding similar results.
Fig. 8. An example of the redshift re-weighting procedure used when comparing the correlation function between galaxies split into bins of galaxy or survey properties. Left: The SkyNet redshift distribution for each half of the NGmix data, split into upper (blue) and lower (red) bins of signal-to-noise ratio \((S/N)\), before (dashed) and after (solid) re-weighting, compared to the full sample \(n(z)\) (black solid curve). Right: The distribution of weights applied to each galaxy to produce the solid \(n(z)\) lines, generated as described in Sec. VI C.

The optimization parameter \(\alpha\), minimizing

\[
||Rv - t|| + ||\alpha(v - I)||
\]

where \(||...||\) denotes the least-squares norm, \(R\) is the matrix of galaxy \(p(z)\)'s each weighted by the lensing weights given in Section II,

\[
R = \begin{bmatrix}
  w_{1p1} & w_{2p1} & w_{3p1} & \ldots & w_{np1} \\
  w_{1p2} & w_{2p2} & w_{3p2} & \ldots & w_{np2} \\
  w_{1p3} & w_{2p3} & w_{3p3} & \ldots & w_{np3} \\
  \ldots & \ldots & \ldots & \ldots & \ldots \\
  w_{1pm1} & w_{2pm2} & w_{3pm3} & \ldots & w_{npmn}
\end{bmatrix}
\]

for \(n\) galaxies and \(m\) photo-z bins with lensing weights \(w_i\) and galaxy \(p(z)\)'s \(p_{ji}\), \(t\) is the target photo-z distribution, \(v\) is the vector of new weights for which we are solving and \(I\) is the identity vector. The parameter \(\alpha\) governs the flexibility of the weight selection — the smaller the value, the better matched the reweighed \(n(z)\) are — and is adjusted to prevent a significant contribution of negative or large weight values, which may impact the validity of the null tests. We find that \(\alpha = 5 \times 10^{-11}\) produces an optimal match between the two half-samples while keeping the weights \(v\) sufficiently regular for our photo-zs and lensing weights. This value may not generalize to other lensing weights or photo-zs. We match the redshift distribution of each half-sample to that of the full sample (i.e. \(t\) is the redshift distribution of the full sample). This procedure is more stable than matching one half to another since smaller weights are needed for each half. The application of the Ridge Regression algorithm then produces a new weight \(v\), which is combined multiplicatively with the lensing weight in the calculation of the correlation functions. The resulting reweighting for galaxies split into bins of low and high galaxy detection signal-to-noise for NGmix is shown in Fig. 8. The left panel shows the \(n(z)\) for each half before (dashed) and after (solid) reweighting compared to the full sample. The corresponding weight histograms are shown in the right panel.

We use the 126 DES SV-shaped mock catalogs described above to compute the variance and significance of the differences between the shear correlation functions in each half-sample. In the mock catalogs, we select a subset of galaxies in narrow redshift slices to match the \(n(z)\) distribution for the full galaxy catalog. Random shape noise is generated from the shear catalog and applied to the mock catalogs, and the property with which we split the galaxy sample in half is then mapped onto the galaxies in each mock via a nearest neighbour algorithm in angular position, and redshift. This preserves the same spatial patterns as exist in the data, but the shears have been randomised so that there is no correlation with this property. We then apply the same procedure to each mock as applied to the data to directly compute the error bars on the difference via Monte Carlo, with the exception of using the true mock point redshift values to reweight the \(n(z)\) histograms of each half instead of a \(p(z)\) estimate for each galaxy. We expect this difference will only underestimate the variance. Any statistically significant deviations then indicate that the there may be a residual systematic error in the shear catalogs related to the quantity split upon, which has affected the measured two-point correlation function.
FIG. 9. Null tests for the \textit{ngmix} two point correlation function based on a variety of catalog and survey properties as described in Table 1. Each panel for a given property shows the difference between the $\xi_+$ relative to its error for the galaxies in the upper and lower halves of the sample split into bins by the magnitude of the quantity. The two halves of the sample have been reweighted to have the same redshift distribution. The error on the difference is computed directly via the mock catalogs. Grey bands are shown representing the 1\$\sigma$ and 2\$\sigma$ variance at each value of $\theta$. Adjacent points in angle are correlated.

2. Results

The split null tests on $\xi_+$ are presented in Figure 9 for \textit{ngmix} and Figure 10 for \textit{im3shape}. This is repeated in Appendix D for $\xi_-$. For each quantity (panel), the difference in $\xi_+$ is shown at each value of $\theta$ relative to the 1\$\sigma$ error in the difference from the mock catalogs. Grey bands corresponding to 1\$\sigma$ and 2\$\sigma$ errors are shown for comparison. The corresponding statistical significance of the null tests for \textit{im3shape} and \textit{ngmix} are given in Table I. We find that for both \textit{ngmix} and \textit{im3shape} the null tests pass with deviations smaller than 2\$\sigma$ ($\chi^2$/d.o.f. = 17.8/8) for all tests except for \textit{ngmix} airmass. Note that because \textit{ngmix} has a higher source density, it is generally more sensitive to residual systematic errors in these tests. While this detection is still weak, it warrants evaluating whether this difference in the galaxy population halves will have a significant bias on the correlation function. To test this, we also show in Table II the difference $\Delta\xi_+ = \xi_+^{(\text{upper})} - \xi_+^{(\text{lower})}$ relative to the 1\$\sigma$ error on the full sample measurement. For \textit{ngmix} airmass, this difference is approximately one-third of the statistical error on the measurements and consistent with the level of bias in several other quantities. Of slightly lesser significance are splits in the magnitude of \textit{ngmix} PSF $e_1$ and $e_2$, for which PSF $e_2$ has the largest difference in $\xi_+$ between upper and lower halves — though still small compared to the statistical error.

There is some subtlety in interpreting the significance of these null tests. First, due to physical effects not accounted for in the simulations, some tests could yield non-zero results but not indicate systematic errors in the data analysis itself. For example, if the level of intrinsic alignments differs between galaxies split by colour, then these null tests could fail and yet the shear measurements themselves could be free of systematic errors. Second, these tests could also, in principle, flag differences between the shear calibrations of galaxies of different types, which although interesting, may not ultimately impact cosmological constraints from the full sample, which could be unbiased on average. Third, as stated above, it is not clear from these tests alone if any
deviations are due to the shear measurements or the photometric redshifts. Finally, note that the $\chi^2$ values from these tests are not independent, due to correlations in the underlying quantities used to construct the tests (e.g., the survey depth is correlated with the seeing). We have performed a large number of null tests, so to the extent that the $\chi^2$ values between many of the tests should be independent, we do expect some apparent deviations purely from statistical fluctuations. However, we have not attempted to combine the tests in order to quote an overall significance.

VII. CONCLUSIONS

In this work, we present cosmic shear two-point measurements from Dark Energy Survey Science Verification data. We find an overall detection significance of $9.7\sigma$ for our higher source density catalog, ngmix. We additionally present multiple advances in band-power estimation, covariance estimation, simulations versus theory, and null tests for shear two-point correlations. Through this work we demonstrate that our measurements are robust and free of statistically significant systematic errors.

We demonstrate that the covariance matrices derived from the DES SV mock simulations presented in this work are consistent with the halo model, including the halo sample variance terms. We also compare the variance in the mock catalogs to the variance in the DES SV data by comparing jackknife covariances computed in the data and mock catalogs. The structure of the covariance matrices is very similar and we detect no statistically significant sources of additional variance in the data.

We find that the B-mode signals in the data are consistent with zero and that the two shear estimation codes agree well. We additionally present a set of simultaneous null tests of the photo-zs and shear measurements, performed by splitting the shear sample in half according to some parameter and comparing the shear correlation functions of the halves. We find that these tests pass with no statistically significant indications of biases. We expect null tests similar to those developed here to have increased utility in future cosmic shear analyses, where the statistical power is larger and the requirements for controlling systematic errors and shear selection effects are more stringent. The DES itself will have nearly $\approx 36\times$ more data and will measure cosmic shear at significantly higher signal-to-noise, so that these tests will be very
TABLE I. Summary of null tests for ngmix and im3shape. Results are given as ngmix (im3shape). The $\chi^2$ values are given for the differences between the two-point correlation function calculated from galaxies that fall within one of two bins in each catalog or survey property. Also shown is the magnitude of the difference relative to the $1\sigma$ error of the measurement of $\xi_+$ on the full sample.

<table>
<thead>
<tr>
<th>Property</th>
<th>$\chi^2$ [d.o.f. = 8]</th>
<th>$\Delta \xi_+ / \sigma(\xi_+)$</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal-to-Noise</td>
<td>4.9 (5.2)</td>
<td>0.05 (0.49)</td>
<td>Signal-to-noise of galaxy detection</td>
</tr>
<tr>
<td>Galaxy Size</td>
<td>5.3 (10.7)</td>
<td>-0.3 (0.15)</td>
<td>Galaxy size (deconvolved with PSF)</td>
</tr>
<tr>
<td>Galaxy Colour</td>
<td>7.3 (2.2)</td>
<td>-0.31 (-0.32)</td>
<td>$g - z$ colour</td>
</tr>
<tr>
<td>Surface Brightness</td>
<td>7.8 (8.7)</td>
<td>0.33 (-0.32)</td>
<td>Galaxy surface brightness</td>
</tr>
<tr>
<td>RA</td>
<td>7.0 (8.8)</td>
<td>0.24 (0.28)</td>
<td>Galaxy right ascension</td>
</tr>
<tr>
<td>Dec</td>
<td>4.0 (6.2)</td>
<td>-0.24 (-0.57)</td>
<td>Galaxy declination</td>
</tr>
<tr>
<td>E(B-V)</td>
<td>5.1 (6.2)</td>
<td>0.23 (0.06)</td>
<td>Mean extinction</td>
</tr>
<tr>
<td>Air Mass</td>
<td>20.7 (13.8)</td>
<td>0.31 (0.46)</td>
<td>Mean $r$-band air mass</td>
</tr>
<tr>
<td>Exposure Time</td>
<td>4.7 (6.8)</td>
<td>0.18 (0.3)</td>
<td>Mean total $r$-band exposure time</td>
</tr>
<tr>
<td>Mag. Limit</td>
<td>4.4 (7.4)</td>
<td>0.18 (0.45)</td>
<td>Mean $r$-band limiting magnitude</td>
</tr>
<tr>
<td>Sky Sigma</td>
<td>1.7 (13.0)</td>
<td>-0.02 (-0.08)</td>
<td>Mean $r$-band RMS sky brightness</td>
</tr>
<tr>
<td>Sky Brightness</td>
<td>5.0 (14.3)</td>
<td>-0.05 (-0.27)</td>
<td>Mean $r$-band sky brightness</td>
</tr>
<tr>
<td>FWHM</td>
<td>6.4 (3.3)</td>
<td>-0.23 (-0.13)</td>
<td>Mean $r$-band PSF FWHM</td>
</tr>
<tr>
<td>PSF $e_1$</td>
<td>16.8 (13.5)</td>
<td>0.12 (-0.37)</td>
<td>Galaxy PSF $e_1$</td>
</tr>
<tr>
<td>PSF $e_2$</td>
<td>17.1 (7.5)</td>
<td>-0.58 (-0.22)</td>
<td>Galaxy PSF $e_2$</td>
</tr>
<tr>
<td>PSF Size</td>
<td>2.6 (5.6)</td>
<td>-0.1 (0.42)</td>
<td>Galaxy PSF size</td>
</tr>
</tbody>
</table>

Future cosmic shear two-point function measurements in the Dark Energy Survey face a variety of challenges. First, while we have a sufficient number of simulations for the SV data, simulating the increased area of the full DES will present a significant computational challenge. This challenge will need to be met by a combination of large simulation campaigns, information compression schemes applied directly to the data, and combinations of theoretical models for the covariances with simulations in order to reduce the noise in the covariance matrix elements. Second, in order to use simulations to evaluate the statistical significance of null tests on future DES data, like those presented in this work, we will need to increase the fidelity of the treatment of both the galaxies and the shear signals. Third, we must better address the formal aspects of the construction of the two-point function statistic estimators in order to make higher precision measurements. Finally, while this work has focused exclusively on broad-bin tomography of the two-point function measurements of cosmic shear, future exploration of higher order correlation functions and finer tomographic binning will be needed to extract the full amount of cosmological information from cosmic shear data. Fortunately, none of these issues are fundamentally intractable and we expect that the new techniques presented in this work will be of great assistance in making future cosmic shear measurements with DES data.

ACKNOWLEDGEMENTS

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

MRB is grateful for the support of the University of Chicago Research Computing Center, and especially Doug Rudd, for the time used to carry out the N-body simulations carried out in this work. MRB would also like to thank Stewart Marshall for his ongoing assistance in using SLAC computing resources. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. JAZ, MAT, SLB acknowledge support from the European Research Council in the form of a Starting Grant with number 240672. MRB and RHW received partial support from NSF-AST-1211838 and from a DOE SciDAC grant. OF and DG were 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’ AA, AR, AN are supported in part by grants 20021_14944 and 20021_1439606 from the Swiss National Foundation. Jarvis has been supported on this project by NSF grants AST-0812790 and AST-1138729. Jarvis, Bernstein, and Jain are partially supported by DoE grant DE-SC0007901. ML is partially supported
Appendix A: Alternative E- and B-mode Statistics

In this appendix we consider alternative statistics of the shear field, verifying that our conclusions above, especially that the B-modes are consistent with zero, do not depend on the choice of statistic. These alternative statistics include the band-powers of Becker and Rozo [46] and power spectra band-powers estimated with PolSpice [16, 17, 48].

1. Band-powers

The band-powers of Becker and Rozo [46] use the methods of Becker [40] to estimate Fourier-space band-powers directly from linear combinations of the real-space two-point functions. The final band-power estimates can be computed from the underlying E-mode power spectrum as

$$E = \int \frac{d\ln \ell^2}{2\pi} C_{EE}(\ell) W_+(\ell)$$

(A1)

where $W_+(\ell)$ is the band-power window function computed from the coefficients $\{f_{+i}, f_{-i}\}$ in Eqs. 3 & 4. See Becker and Rozo [46] for more details. The optimal computation of the band-powers requires computing the effective radial bin window functions of the shear correlation function points. Instead in this work we just use the geometric approximation to the bin window functions to compute the amplitudes $\{f_{+i}, f_{-i}\}$. This procedure means that the band-powers do not separate E- and B-modes as well as they could in principle. However, when comparing to a fiducial cosmological model below, we do compute the band-power window function using estimates of radial bin window functions from the data. These window functions are computed via interpolating the weighted counts in each radial bin of the estimated shear two-point function. We have compared the results of this procedure for computing the window functions to estimates of the window functions from counts in finer bins. We find unsurprisingly that the bin window functions are quite smooth and thus the interpolation is accurate enough for our purpose.

2. Spherical Harmonic Power Spectrum

The cosmic shear power spectrum can also be estimated in spherical harmonic space, which has the advantage of being faster and less memory intensive than working in real-space. In view of upcoming wide field galaxy lensing surveys, e.g. the full five year DES dataset, we therefore investigate the applicability of standard spherical harmonic space methods to weak lensing. For this purpose, we use the PolSpice [17, 48] code together with the HEALPix [68] package, which has been applied to, amongst other things, CMB polarization data [e.g., 69].

16 http://www2.iap.fr/users/hivon/software/PolSpice/
show the band-power window functions for the shape noise appropriate for each catalog. The dotted lines are calculated from the simulation realizations using the model given above. Note that the theoretical prediction has been convolved with the Gaussian window of \(22 \text{arcmin}^2\). The solid line is the prediction for the shear power spectrum for the flat, \(\Lambda\)CDM model given above. The dashed line shows the integral of the band-power window functions over the shear power spectrum.

For our analysis, we pixelize the galaxy ellipticities onto a HEALPix pixelization of the sphere with a resolution of \(N_{\text{side}}=1024\), where each pixel covers a solid angle of \(\approx 11.8 \text{arcmin}^2\). In order to obtain a robust estimate of the shear field, we need to correct for multiplicative bias in the measured ellipticities. Since the correction factors described in Sections II A and II B are noisy estimates of the true corrections, we determine the mean sensitivity or multiplicative bias correction for our galaxy sample and apply this mean correction to the pixelized maps. As the power spectrum is estimated from maps constructed from the discrete values of the galaxy ellipticities, we apply a conservative masking scheme to maximize galaxy number density. We therefore adopt the DES SV LSS mask used for galaxy clustering measurements [70]. This mask is identical to the DES SV mask used for weak lensing except that it restricts analyses to the largest contiguous region overlapping the SPT-E field by selecting the area with \(60 < \text{ra [deg]} < 95\) and \(-60 < \text{dec [deg]} < -40\). It further considers only regions with survey limiting magnitude in the i-band > 22.5 (i.e. all regions considered to provide at least 10\(\sigma\) measurements for objects at i-band = 22.5; [70]). For the power spectrum measurement, we limit all integrations to scales smaller than \(\theta_{\text{max}} = 15\) degrees and we apodize the correlation function with a Gaussian window of \(\theta_{\text{FWHM}} = 10\) degrees. Finally, we correct the measured power spectra for the HEALPix pixel window function and compress them into 7 band-powers with PolSpice band-power kernels.

The noise power spectrum needs to be computed from

FIG. 11. Band-powers in DES SV data for ngmix (top) and im3shape (bottom). The error bars indicated by the grey bands are calculated from the simulation realizations using the shape noise appropriate for each catalog. The dotted lines show the band-power window functions \(W_+ (\ell)\) scaled so that their peak values are \(2 \times 10^{-6}\). The solid line is the prediction for the shear power spectrum for the flat, \(\Lambda\)CDM model given above. The dashed line shows the integral of the band-power window functions over the shear power spectrum.

FIG. 12. Spherical harmonic shear power spectrum estimated using PolSpice. The left and right panels correspond to the ngmix and im3shape catalogs, respectively. The top and bottom panels show the E- and B-modes, respectively. The measurement uncertainties are estimated using the mock catalogs. The black solid lines show the predictions for the flat, \(\Lambda\)CDM model given above. Note that the theoretical prediction has been convolved with the PolSpice kernels, which relate the true to measured power spectra. The \(S/N\) values for the E-modes are computed as outlined in Section IV A and the \(\chi^2\) values for the B-modes indicate consistency with zero. The reported values take into account correlations between the band-powers.
simulations. In order to produce noise only maps from the data, we remove correlations in the ellipticity maps by rotating each galaxy ellipticity by a random angle. We then estimate the noise power spectrum as the mean of the power spectra of 100 such random realizations. This procedure yields shape noise estimates consistent with \( C_{\ell,SN} = \sigma_{\epsilon,\text{pix}}^2 \) where \( \sigma_{\epsilon,\text{pix}}^2 \) is the variance of either component of the mean ellipticity per pixel and \( n_{\text{pix}} \) is the angular number density of HEALPix pixels. Comparing the measured shape noise to the galaxy-based Gaussian shape noise estimate \( C_{\ell,SN} = \sigma_{\epsilon,\text{gal}}^2 \), where \( \sigma_{\epsilon,\text{gal}}^2 \) is the variance of either component of the galaxy ellipticities and \( n_{\text{gal}} \) denotes the galaxy number density, we find that the latter underestimates the measured shape noise. This suggests that the galaxy ellipticity distribution is non-Gaussian and the Gaussian approximation can therefore only be applied after averaging the galaxy ellipticities over pixels. We test the pipeline using Gaussian field realizations and the mock catalogs.

3. Results

Figure 11 shows the non-tomographic band-powers using the methods of Becker and Rozo [46], their window functions as the dotted lines, and their error bars computed with the mock catalogs as the grey bands. We find a detection significance 6.1\( \sigma \) and 5.7\( \sigma \) for NGMIX and IM3SHAPE, respectively. These detection significances are similar to the real-space two-point functions. Finally, the solid line shows the expected shear power spectrum amplitude assuming the flat, \( \Lambda \)CDM model given above. The dashed line shows for each band-power the integral of the band-power window function over the shear power spectrum.

Figure 12 shows the results for the PolSpice statistics. We find a detection of cosmic shear of 5.6\( \sigma \) and 5.4\( \sigma \) for NGMIX and IM3SHAPE respectively for the PolSpice statistics. Note that the PolSpice statistics do not use as many high-\( \ell \) modes as the real-space band-powers or the real-space correlation functions, so that one expects a lower detection significance. We also find that the B-modes are statistically consistent with zero for the PolSpice statistics.

Finally, note that these two estimators process the data in different ways (e.g., averaging the data in pixels versus computing real-space correlation functions), have different sensitivities to shot noise, and have different Fourier-space window functions. We thus do not expect them to give precisely the same results in Fourier-space for the shear power spectrum. However, we do expect that when treated self-consistently they should give statistically consistent results for cosmological parameters, as demonstrated in the accompanying cosmological analysis of this data [1].

Appendix B: Validation of the Mock Catalogs

In this section we present a validation test for the mock catalogs. We compare the shear correlation functions measured in the mock catalogs in tomographic bins with the theoretical expectation from the Takahashi et al. [14] fitting function for the matter power spectrum. The result of this test is shown in Figure 13. We find that at high redshift the small-scale shear correlation functions are suppressed relative to the theoretical expectation. Note however that this numerical effect is below the scales where the two-point functions are being used for cosmological parameter estimation (\( \approx 2 - 4 \) arcminutes for \( \xi_+ \) and \( \approx 25 - 55 \) arcminutes for \( \xi_- \); see Table 2 of The Dark Energy Survey Collaboration et al. [1]). Additionally, we only estimate the covariance of the two-point functions from the mock catalogs, not the mean signal. Within the noise of our mock covariance matrix, the overall parameter uncertainties are consistent when using the halo model versus the simulation covariance (see Sec. 4 for a quantitative comparison). This fact may indicate that the covariance is less sensitive to these numerical effects than the mean signal. Future work may require higher-resolution shear fields for covariance estimation.

Appendix C: Detailed Covariance Matrix Validation

In this section, we present further details of the validation of the covariance matrices, including our tomographic halo model computations and the comparison to the simulations. The halo model covariance was computed with the CosmoLike covariance module (see Eifler et al. [49] and Krause et al. [50] for details).

In the halo model, the covariance of tomographic shear power spectra \( C^{ij}_\kappa(l) \) is given by [71,73]

\[
\text{Cov} \left( C^{ij}_\kappa(l_1), C^{kl}_\kappa(l_2) \right) = \frac{2\pi \delta_{l_1,l_2}}{\Omega_{\Delta l_1} \Delta l_1} \left[ \left( C^{ij}_\kappa(l_1) + \delta_{il} \sigma_{\epsilon}^2 \right) \left( C^{kl}_\kappa(l_2) + \delta_{jl} \sigma_{\epsilon}^2 \right) + \left( C^{il}_\kappa(l_1) + \delta_{il} \sigma_{\epsilon}^2 \right) \left( C^{jk}_\kappa(l_2) + \delta_{jk} \sigma_{\epsilon}^2 \right) \right] + \int_{|l| \leq l_1} \frac{d^3l}{A(l_1)} \int_{|\nu| \leq l_2} \frac{d^3\nu}{A(l_2)} \left[ \frac{1}{\Omega_n} T^{ij,kl}_{\kappa,0}(\nu, l, -\nu, -l) + T^{ij,kl}_{\kappa,\text{HSV}}(\nu, l, -\nu, -l) \right],
\]

with \( n' \) the number of source galaxies in tomography bin \( i \), \( \sigma_{\epsilon} \) the ellipticity dispersion, \( A(l_i) = \int_{|l| \leq l_i} d^3l \approx \frac{2\pi \delta_{l_1,l_2}}{\Omega_{\Delta l_1} \Delta l_1} \left[ \left( C^{ij}_\kappa(l_1) + \delta_{il} \sigma_{\epsilon}^2 \right) \left( C^{kl}_\kappa(l_2) + \delta_{jl} \sigma_{\epsilon}^2 \right) + \left( C^{il}_\kappa(l_1) + \delta_{il} \sigma_{\epsilon}^2 \right) \left( C^{jk}_\kappa(l_2) + \delta_{jk} \sigma_{\epsilon}^2 \right) \right] + \int_{|l| \leq l_1} \frac{d^3l}{A(l_1)} \int_{|\nu| \leq l_2} \frac{d^3\nu}{A(l_2)} \left[ \frac{1}{\Omega_n} T^{ij,kl}_{\kappa,0}(\nu, l, -\nu, -l) + T^{ij,kl}_{\kappa,\text{HSV}}(\nu, l, -\nu, -l) \right],
\]

\[
\text{Cov} \left( C^{ij}_\kappa(l_1), C^{kl}_\kappa(l_2) \right) = \frac{2\pi \delta_{l_1,l_2}}{\Omega_{\Delta l_1} \Delta l_1} \left[ \left( C^{ij}_\kappa(l_1) + \delta_{il} \sigma_{\epsilon}^2 \right) \left( C^{kl}_\kappa(l_2) + \delta_{jl} \sigma_{\epsilon}^2 \right) + \left( C^{il}_\kappa(l_1) + \delta_{il} \sigma_{\epsilon}^2 \right) \left( C^{jk}_\kappa(l_2) + \delta_{jk} \sigma_{\epsilon}^2 \right) \right] + \int_{|l| \leq l_1} \frac{d^3l}{A(l_1)} \int_{|\nu| \leq l_2} \frac{d^3\nu}{A(l_2)} \left[ \frac{1}{\Omega_n} T^{ij,kl}_{\kappa,0}(\nu, l, -\nu, -l) + T^{ij,kl}_{\kappa,\text{HSV}}(\nu, l, -\nu, -l) \right],
\]
FIG. 13. The shear correlation functions in the mock catalogs compared to the expected values from Takahashi et al. [45] for all three tomographic bins (labeled in the top left corner from left to right). In the top panel, the solid lines show the theoretical expectation, the bands show the 1σ sample variance estimate and the dashed line shows the mean from the mock catalogs. ξ+ is in red and ξ− is in blue. In the bottom panels, we show the fractional deviation of the mean signal in the mock catalogs from the expected values from Takahashi et al. [45] in units of the sample variance. ξ+ data below ≈ 2−4 arcminutes and ξ− data below ≈ 25−55 arcminutes is not used for the final cosmological analysis in The Dark Energy Survey Collaboration et al. [1] due to the expected baryonic effects in the matter power spectrum.

\[2\pi l_i \Delta l_i, \text{the integration area associated with a power spectrum bin centered at } l_i \text{ and width } \Delta l_i, \text{ and } T_{ijkl}^{ijkl} \text{ and } T_{\kappa,0}^{ijkl} \text{ the convergence trispectrum of source redshift bins } i, j, k \text{ and } l \text{ in the absence of finite volume effects and the halo sample variance contribution to the trispectrum [51, 73]. Our halo model implementation for these terms is described in Eifler et al. [74].}

Note that Equation (C1) ignores the so called finite-area effect (cf. Sato et al. [75] or Friedrich et al. [54]), linear beat-coupling terms [e.g., 76] and linear dilation terms Li et al. [e.g., 77]. For a survey of the size of DES-SV the finite-area effect is expected to be negligible. Furthermore, ignoring this effect is at most conservative since it will slightly overestimate the statistical uncertainties. The beat-coupling terms are negligible compared to the halo sample variance terms (and even the non-Gaussian terms, see e.g., Takada and Jain [76]). Further, the linear dilation terms reduce the effect of the beat-coupling terms and are negligible [77]. Finally, we have ignored the effects of masking (except for the total area of the survey in the halo sample variance terms). We have found with Gaussian simulations that the effects of the details of the mask, besides the overall survey area, are negligible when computing cosmological constraints.

The covariance of angular shear correlation functions is then given by

\[\text{Cov} \left( \xi_{ij}^{+}(\theta_1), \xi_{kl}^{+}(\theta_2) \right) = \int \frac{dl}{2\pi} J_{0/4}(l\theta_1) \int \frac{dl'}{2\pi} J_{0/4}(l'\theta_2) \text{ Cov} \left( C_{ij}^{+}(l_1), C_{kl}^{+}(l_2) \right) \]  

(C2)

where we use the results of Joachimi et al. [78] to simplify the calculation of the Gaussian part of the covariance.

Figure 14 shows the full tomographic correlation matrix, comparing the halo model on the lower-right and
FIG. 14. Comparison of the shear correlation function correlation matrix estimated from mock catalogs and calculated from the halo model. Figure 4 shows a subset, those for tomographic bin combinations (1,1), (1,3) and (3,3), of the covariance matrix elements shown in this figure. The correlation matrix from mock catalogs is on the upper-left and that from the halo model is on the lower-right.

the mock catalogs on the upper-left. The overall structure of the covariance matrices is similar in both computations, but the mock catalogs exhibit more noise in the off-diagonal components.

Appendix D: Additional Two-Point Null Tests of \( \xi^- \)

We have repeated an identical analysis for \( \xi^- \) to that described in Sec. VTC for \( \xi^+ \). We show the results of the tests for IM3SHAPE in Fig. 15 and for NGMIX in Fig. 16. Qualitatively, comparing to Figs. 9 & 10, there is an indication that some of the larger deviations in the figures for \( \xi^+ \) may be due to additive systematic errors. For example, there is an offset in the difference of \( \xi^+ \) based on values of airmass at the 2\( \sigma \) level that disappears for \( \xi^- \). The corresponding \( \chi^2 \) and difference values are given in Table II. There are no significant indications of systematic errors in these null tests for \( \xi^- \), though this may simply be due to the poorer constraining power of \( \xi^- \).

FIG. 15. Null tests for the ngmix two point correlation function based on a variety of catalog and survey properties as described in Table 1. See Fig. 9 for details.


FIG. 16. Null tests for the im3shape two point correlation function based on a variety of catalog and survey properties as described in Table 1. See Fig. 9 for details.


TABLE II. Summary of null tests for \textit{ngmix} and \textit{im3shape}.

The $\chi^2$ of two bins in each catalog or survey property. Also shown is the magnitude of the difference relative to the 1σ error of the measurement of $\xi_\pm$ on the full sample.

<table>
<thead>
<tr>
<th>Property</th>
<th>$\chi^2$ [d.o.f. = 8]</th>
<th>$\Delta \xi_\pm/\sigma(\xi_\pm)$</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal-to-Noise</td>
<td>5.8 (1.8)</td>
<td>-0.07 (0.03)</td>
<td>Signal-to-noise of galaxy detection</td>
</tr>
<tr>
<td>Galaxy Size</td>
<td>2.5 (5.0)</td>
<td>-0.23 (-0.35)</td>
<td>Galaxy size (deconvolved with PSF)</td>
</tr>
<tr>
<td>Galaxy Colour</td>
<td>7.1 (3.8)</td>
<td>-0.3 (0.04)</td>
<td>$g - z$ colour</td>
</tr>
<tr>
<td>Surface Brightness</td>
<td>4.4 (5.2)</td>
<td>-0.04 (-0.06)</td>
<td>Galaxy surface brightness</td>
</tr>
<tr>
<td>RA</td>
<td>2.9 (3.0)</td>
<td>0.06 (-0.22)</td>
<td>Galaxy right ascension</td>
</tr>
<tr>
<td>Dec</td>
<td>4.9 (3.5)</td>
<td>-0.35 (-0.37)</td>
<td>Galaxy declination</td>
</tr>
<tr>
<td>E(B-V)</td>
<td>2.8 (4.9)</td>
<td>-0.22 (-0.02)</td>
<td>Mean extinction</td>
</tr>
<tr>
<td>Air Mass</td>
<td>2.7 (3.4)</td>
<td>-0.01 (-0.08)</td>
<td>Mean $r$-band air mass</td>
</tr>
<tr>
<td>Exposure Time</td>
<td>4.5 (2.5)</td>
<td>-0.35 (0.0)</td>
<td>Mean total $r$-band exposure time</td>
</tr>
<tr>
<td>Mag. Limit</td>
<td>2.2 (3.3)</td>
<td>-0.29 (-0.43)</td>
<td>Mean $r$-band limiting magnitude</td>
</tr>
<tr>
<td>Sky Sigma</td>
<td>3.8 (5.6)</td>
<td>-0.21 (-0.3)</td>
<td>Mean $r$-band RMS sky brightness</td>
</tr>
<tr>
<td>Sky Brightness</td>
<td>4.0 (6.2)</td>
<td>-0.27 (-0.42)</td>
<td>Mean $r$-band sky brightness</td>
</tr>
<tr>
<td>FWHM</td>
<td>4.1 (4.5)</td>
<td>-0.2 (-0.08)</td>
<td>Mean $r$-band PSF FWHM</td>
</tr>
<tr>
<td>PSF $e_1$</td>
<td>2.7 (7.9)</td>
<td>-0.37 (-0.55)</td>
<td>Galaxy PSF $e_1$</td>
</tr>
<tr>
<td>PSF $e_2$</td>
<td>6.8 (5.8)</td>
<td>-0.5 (-0.33)</td>
<td>Galaxy PSF $e_2$</td>
</tr>
<tr>
<td>PSF Size</td>
<td>1.2 (3.8)</td>
<td>-0.08 (-0.1)</td>
<td>Galaxy PSF size</td>
</tr>
</tbody>
</table>


