Dark Energy Survey Year 3 results: Cosmology with moments of weak lensing mass maps

M. Gatti, B. Jain, C. Chang, M. Raveri, D. Zürcher, L. Secco, L. Whiteway, N. Jeffrey, C. Doux, T. Kacprzak, D. Bacon, P. Fosalba, A. Alarcon, A. Amon, K. Bechtol, M. Becker, G. Bernstein, J. Blazek, A. Campos, A. Choi, C. Davis, J. Derose, S. Dodelson, F. Elsner, J. Elvin-Poole, S. Everett, A. Ferte, D. Gruen, I. Harrison, D. Huterer, M. Jarvis, E. Krause, P. F. Leget, P. Lemos, N. Maccrann, J. McCullough, J. Muir, J. Myles, A. Navarro, S. Pandey, J. Prat, R. P. Rolls, A. Roodman, C. Sanchez, E. Sheldon, M. Troxel, I. Tutusaus, B. Yin, M. Aguena, S. Allam, F. Andrade-Oliveira, J. Annis, E. Bertin, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, R. Cawthon, M. Costanzi, L. da Costa, J. De Vicente, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, D. L. Hollowood, K. Honscheid, D. J. James, M. A. G. Maia, J. L. Marshall, P. Melchior, F. Menanteau, M. Miquel, R. Morgan, A. Palmese, A. Pieres, M. Schubnell, S. Serrano, I. Sevilla-Noarbe, and T. N. Varga

(DES Collaboration)

1Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
2Department of Astronomy and Astrophysics, University of Chicago, Chicago, Illinois 60637, USA
3Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA
4Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland
5Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom
6Laboratoire de Physique de l’Ecole Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, Paris 75005, France
7Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, United Kingdom
8Institut d’Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
9Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
10Argonne National Laboratory, 9700 South Cass Avenue, Lemont, Illinois 60439, USA
11Kavli Institute for Particle Astrophysics and Cosmology, P. O. Box 2450, Stanford University, Stanford, California 94305, USA
12Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, Wisconsin 53706-1390, USA
13Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
14Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA
15Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
16Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15312, USA
17California Institute of Technology, 1200 East California Blvd, MC 249-17, Pasadena, California 91125, USA
18Kavli Institute for Particle Astrophysics and Cosmology, P. O. Box 2450, Stanford University, Stanford, California 94305, USA
19Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA
20Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15312, USA
21NSF AI Planning Institute for Physics of the Future, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
22Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, Ohio 43210, USA
I. INTRODUCTION

Gravitational lensing is one of the cleanest probes for studying the mass distribution in the Universe. General relativity predicts that the trajectories of photons emitted by distant galaxies are bent as they pass through regions of space-time perturbed by the mass distribution between the galaxy and the observer [1]. When studying the light emitted by distant galaxies, the level of distortion induced by the mass distribution of the Universe, or large scale structure (LSS), is usually small, at the percent level—the regime of weak gravitational lensing. By collecting observations and measuring the shapes of many galaxies, statistical tools can be used to infer the mass distribution of the Universe [2–8]. Ongoing and future surveys (DES, Dark Energy Survey Collaboration [9]; Kilo-Degree Survey KIDS, Kuijken et al. [10]; Hyper Suprime-Cam HSC, Aihara et al. [11]; Vera C. Rubin Observatory’s Legacy Survey, LSST Science Collaboration et al. [12]; Euclid, Laureijs et al. [13]) are currently measuring (or planning to measure) the shapes of tens to hundreds of millions of galaxies, spanning thousands of square degrees of the sky. In particular, DES recently measured 100 million galaxies spanning ∼5000 square degrees of the southern hemisphere [14], and created the largest map of the mass distribution of the Universe from a galaxy survey [4].

For a given cosmological model, the statistical properties of the mass distribution can be predicted over time. Second-order statistics, such as correlation functions [15–19], the power spectrum [20], or the waveletlike COSEBIs...
(complete orthogonal sets of E/B integrals) [21], are standard tools used to exploit the Gaussian information of the mass maps. However, a weak lensing mass map contains information beyond that captured by second-order statistics, as its probability distribution function (PDF) has non-Gaussian features induced by gravitational evolution. In particular, the PDF of the mass distribution in the late Universe is roughly approximated by a log-normal [22–24], a fact that has also been investigated for the weak lensing convergence field with DES data [25].

Higher-order statistics are appealing, as their use can improve constraints on cosmological parameters [26–29] over standard 2-point statistics, or can help discriminate between extended models such as modified gravity theories [30,31]. Numerous tools have been developed to extract the non-Gaussian information from mass maps. Higher-order statistics commonly used with weak lensing include shear peak statistics [5,29,31–37], higher moments of the weak lensing convergence field [2,7,26,27,31,38,39], three-point correlation functions or bispectra [40–43], Minkowski functionals [27,44–46], and machine-learning methods [47–50]. Many of these have recently been applied to data [5,34,36,47,49], often performing well in terms of cosmological constraints. The theoretical modeling of some of these statistics is often complex, and large suites of N-body simulations, spanning the parameter space considered in the analysis, are used to model the observables.

This work focuses on the use of second and third moments of weak lensing mass maps to constrain cosmology. Moments have been studied in the past, and have been measured both in data and simulations [2,7,26,27,51–54], although they have not been used to place constraints on cosmological parameters. Tests using simulations have shown improvements to cosmological constraints arising from using moments of order higher than second [26–28]. The methodology used in this paper has been developed and tested using simulations in a companion paper, [26] (hereafter G20). Although the methodology can be applied to any dataset, the analysis in G20 was geared towards the first three years of data of DES. The modeling of second and third moments developed in G20 is based on theoretical predictions, therefore it does not rely on large suites of N-body simulations (though the predictions are tested against simulations); moreover, observational systematics errors such as photometric redshift uncertainties or intrinsic alignment are modeled and marginalized during the analysis. This work applies that methodology to the first three years of data (Y3) from DES, presenting the cosmological constraints, discussing a number of observational systematic null tests, and comparing the results with constraints from other DES Y3 probes and/or external datasets (e.g., Planck).

The paper is organized as follows: Section II describes the data and simulations used in this work; Sec. III provides a short description of the theoretical modeling of the observables used in the analysis (the second and third moments of the convergence field); Sec. IV describes the likelihood and the covariance used in the cosmological parameter inference, and discusses the priors adopted in the analysis; Sec. V summarizes the preunblinding tests; Sec. VI presents the cosmological results, along with a number of internal consistency tests and comparisons with results from other DES analyses or from analyses using external data sets; Sec. VII summarizes our findings.

II. DATA AND SIMULATIONS

A. Data

The main goal of our analysis is to measure second and third moments of the convergence field and use them to estimate cosmological parameters. To this aim, we use the weak lensing catalog from the first three years (Y3) of the DES [14].

DES [9] is a six-year survey that spans ~5000 deg² of the southern hemisphere. Images have been taken in grizY filters by the 570 megapixel Dark Energy Camera [DECam, [55]], mounted on the Cerro Tololo Inter-American Observatory (CTIO) four-meter Blanco telescope in Chile. The raw images were processed by the DES Data Management (DESDM) team [56–58]. Full details about the image processing are provided in [56,57].

The DES Y3 weak lensing sample is described in detail in [14] and builds upon the Y3 Gold catalog [59]. It is created using the METACALIBRATION algorithm [60,61], which infers the galaxy ellipticities starting from noisy images of the detected objects in the r, i, z bands. The METACALIBRATION algorithm was used previously in the DES Y1 analysis [62]. METACALIBRATION uses an approximate estimator of the shear field and self-calibrates it using the response of the estimator to shear as well as to selection effects. A number of selection cuts are designed to remove objects in the catalog potentially affected by systematic effects [14]. An inverse variance weight is also assigned to galaxies in order to enhance the overall signal-to-noise. The final DES Y3 shear catalog has 100,204,026 objects, with a weighted n_eff = 5.59 galaxies arcmin⁻², over an effective area of 4139 square degrees.

Although the METACALIBRATION self-calibration procedure removes most of the multiplicative bias, for the DES Y3 weak lensing sample there is a known residual additional multiplicative bias at the level of 2 or 3 per cent [63]. This residual bias stems mostly from a shear-redshift-dependent detection bias due to blending of galaxy images, for which the METACALIBRATION implementation adopted in DES Y3 is unable to account [64]. We do not calibrate for this factor at the catalog level, but we do marginalize over it in the analysis. In [14] the weak lensing sample has also been tested for additive biases (e.g., due to point-spread-function residuals). In particular, the catalog is characterized by a nonzero mean shear whose origin is...
unknown and which is subtracted at the catalog level before performing any analysis.

The weak lensing sample is divided into four tomographic bins of roughly equal number density using the SOMPZ method [65]; SOMPZ, in combination with constraints from clustering redshifts [66], also provides redshift distribution estimates (see Fig. 1). The $n(z)$’s are further tweaked to take into account the redshift-dependent effects of blending [63]. During the cosmological analysis, additional constraints on the redshift distributions are provided by shear ratios [67]. Shear ratios are ratios of small-scale galaxy-galaxy lensing measurements obtained using different source samples (in this case, different weak lensing tomographic bins) and a common lens sample. Not only do they improve constraints on redshift distributions, but they also help constraining both intrinsic alignment parameters and cosmological parameters.

A two-stage blinding scheme was implemented for all DES Y3 cosmological analyses in order to avoid intentional or unintentional confirmation bias. First, the weak lensing sample was blinded by means of a multiplicative factor, in a fashion similar to what was adopted in the Y1 analysis [62]. After all the catalog and map-based systematic tests were passed [4, 14], the hidden value was revealed at the second level of blinding, which follows the work of [68], the catalog had already been validated and unblinded. The procedure allows us to generate 1000 independent simulated shear catalogs.

\begin{equation}
\Delta_i^{\text{blinded}} = \hat{d}_i + d_i(\Theta_{\text{ref}} + \Delta\Theta) - d_i(\Theta_{\text{ref}}).
\end{equation}

In the above equation, $d_i$ is a theory data vector computed at a given cosmology $\Theta$; $\Theta_{\text{ref}}$ is a fiducial cosmology (we used the DES Y1 $3 \times 2\text{pt}$ cosmology from [69]), and $\Delta\Theta$ is a blind shift in the cosmological parameters (drawn from a distribution three times larger than the DES Y1 $3 \times 2\text{pt}$ posterior).

A number of systematic tests were performed on blinded data vectors (see Sec. V) before proceeding to inspect the unblinded cosmological results.

## B. Simulations

Covariance matrices for our measurement are generated: (a) for our fiducial covariance, using log-normal realizations from FLASK [70], (b) for testing, using the N-body simulation hereafter called “T17” [71], and (c) also for testing, using the N-body simulation PKDGRAV [72]. Moreover, both T17 and PKDGRAV simulations are used to validate our modeling (Appendix B). Such a validation was already performed in G20, but using only T17 simulations; we repeat that here, for both sets of N-body simulations, with updated analysis choices.

### 1. FLASK realizations

We use the FLASK (Full-sky Lognormal Astro-fields Simulation Kit) software [70] to rapidly generate full-sky, log-normal realizations of the convergence field. FLASK assumes the convergence field to be described by a zero-mean shifted log-normal distribution, where the parameters of the log-normal probability distribution function (PDF) are chosen to match the variance and skewness of the input. We use here the 1000 independent FLASK realizations produced for the validation of the DES Y3 $3 \times 2\text{pt}$ covariance [73]. The lognormal approximation for the covariance has been shown to be sufficient to not bias the recovery of the cosmological parameters in G20. The cosmological parameters of the input power spectra used for the FLASK realizations are $\Omega_m = 0.3$, $\sigma_8 = 0.82355$, $\Omega_b = 0.048$, $n_s = 0.97$, $h_{100} = 0.69$ and $\Omega_c h = 0.0083$. We also assumed DES Y3 redshift distributions. The FLASK convergence realizations were provided in maps using the Hierarchical Equal Area isoLatitude Pixelation scheme (HEALPix, [74]) with resolution NSIDE = 4096. In order to create a simulated weak lensing galaxy catalog, we then used the position, shape noise (obtained by randomly rotating each galaxy), and weight of the galaxies of the fiducial DES Y3 weak lensing catalog; depending on the position of each individual galaxy, we sampled the simulated shear maps and added shape noise accordingly. This procedure allows us to generate 1000 independent simulated shear catalogs.

### 2. T17 simulations

The first set of N-body simulations used in this work are the T17 [71] simulations. The set consists of 108 full-sky lensing convergence and shear maps, spanning a wide redshift range (between $z = 0.05$ and 5.3) at intervals of 150 $h^{-1}$ Mpc comoving distance. The N-body simulations...
assume a WMAP 9 cosmology ($\Omega_{m} = 0.279$, $\sigma_{8} = 0.82$, $\Omega_{b} = 0.046$, $n_{s} = 0.97$, $h = 0.7$), and were run using L-GADGET2 [75]. Initial conditions were generated using 2LPTIC [76]. The simulations begin with 14 boxes in steps of 450 $h^{-1}$ Mpc, with total side lengths of $L = 450, 900, 1350, \ldots, 6300$ $h^{-1}$ Mpc. There are six independent copies at each box size and 2048$^{3}$ particles per box. Lens plane snapshots are taken at intervals of 150 $h^{-1}$ Mpc comoving distance. The expected accuracy of the average matter power spectra from the simulations (compared to predictions from the revised HALOFIT, [77]) is within 5% for $k < 1$ $h$ Mpc$^{-1}$ at $z < 1$, for $k < 0.8$ $h$ Mpc$^{-1}$ at $z < 3$, and for $k < 0.5$ $h$ Mpc$^{-1}$ at $z < 7$ [71]. Weak lensing quantities for each simulation were estimated using the multiple plane ray-tracing algorithm GRAYTRIX [78], and shear and convergence HEALPIX maps with resolution $\text{NSIDE} = 4096$ are provided.

For each of the 108 simulations, we cut out four independent (i.e., nonoverlapping) regions corresponding to the DES Y3 footprint. We then stacked the convergence and shear snapshots at different redshift to produce convergence and shear maps for the four weak lensing tomographic bins. This gave us 432 independent realizations of the shear field for each tomographic bin. In order to create a simulated weak lensing galaxy catalog, we used the position, shape noise (obtained by randomly rotating each galaxy), and weight of the galaxies of the fiducial DES Y3 weak lensing catalog; depending on the position of each individual galaxy, we sampled the simulated shear maps and added shape noise accordingly. We ended up with 432 independent simulated shear catalogs from T17 simulations.

3. PKDGRAV simulations

The second set of N-body simulations is the DarkGridV1 suite, produced using the PKDGRAV3 code [72] and described in detail in [29,79]. In particular, we use 50 independent realizations at the fixed cosmology $\Omega_{m} = 0.26$, $\sigma_{8} = 0.84$, $\Omega_{b} = 0.0493$, $n_{s} = 0.9649$, $h = 0.673$. All simulations include three massive neutrino species with a mass of $m_{\nu} = 0.02$ eV per species [79]. The simulations were obtained using 14 replicated boxes in each direction ($14^{3}$ replicas in total) so as to span the redshift interval between $z = 0$ and $z = 3$. Each individual box contains 768$^{3}$ particles and has a side length of 900 $h^{-1}$ Mpc. Such a configuration is known to yield a field variance that is too small at very large scales [47]; however, such scales are not considered in this work. For each simulation, lens planes are provided at ~87 redshifts between $z = 3.0$ and $z = 0.0$, equally spaced in proper time. Lensing quantities (shear and convergence) were obtained under the Born approximation. For each simulation, we cut out four independent DES Y3 footprints and thereby created 200 independent catalogs in a fashion similar to the T17 simulations.

III. THEORETICAL MODELING

We provide here a short summary of the theoretical modeling of our observables. Further details are provided in G20.

Our cosmological analysis relies on the theoretical modeling of the second and third moments of convergence maps, which is based on cosmological perturbation theory [80–82]. Consider three convergence maps, obtained from different tomographic bins (labeled $i, j, k$) of the weak lensing catalog (the equations below apply for more tomographic bins as well, taken two or three at a time). The maps are smoothed by a top-hat filter of smoothing length $\theta_{0}$. The second and third moments are then given by

$$\langle \kappa^{2} \rangle_{\theta_{0}}^{i,j,EE/BB} = \int d\chi \frac{q(\chi)q(\chi)}{\chi^{2}} \times \sum_{\ell} \frac{2\ell + 1}{4\pi} f_{\ell} W_{\ell}(\theta_{0})^{2} \times \sum_{\ell'} M_{\ell'\ell}^{EE/BB,EE} P_{NL}(\ell' / \chi, \chi) F_{\ell'} F_{\ell},$$

$$\langle \kappa^{3} \rangle_{\theta_{0}}^{i,j,k,EE/BB} = \int d\chi \frac{q(\chi)q(\chi)q(\chi)}{\chi^{4}} \times S_{3} \left[ \sum_{\ell} \frac{2\ell + 1}{4\pi} f_{\ell} W_{\ell}(\theta_{0})^{2} \times \sum_{\ell'} M_{\ell'\ell}^{EE/BB,EE} P_{NL}(\ell' / \chi, \chi) F_{\ell'} F_{\ell} \right]^{2}.$$  

Here the lensing kernel term $q(\chi)$ is given by

$$q(\chi) = \frac{3H^{2}_{0} \Omega_{m}}{2c^{2}} \frac{\chi}{a(\chi)} \int_{\chi}^{z_{s}} dz' n'(z') dz' \frac{\chi' - \chi}{\chi}.$$  

where $\chi$ is comoving distance, $\chi_{h}$ is the horizon comoving distance, $H_{0}$ the Hubble constant at the present time, $c$ the speed of light, $n'(z)$ the normalized redshift distribution of a given tomographic bin, and $a(\chi)$ the scale factor. Furthermore, in Eqs. (2) and (3), $W_{\ell}(\theta_{0})$ represents the top-hat filter of smoothing length $\theta_{0}$ in harmonic space, defined as

$$W_{\ell}(\theta_{0}) = \frac{P_{\ell-1}(\cos(\theta_{0})) - P_{\ell+1}(\cos(\theta_{0}))}{(2\ell + 1)(1 - \cos(\theta_{0}))},$$

where $P_{\ell}$ are Legendre polynomials of order $\ell$. Other terms in Eqs. (2) and (3) are the mode-coupling matrices $M_{\ell\ell'\ell''\ell'}^{EE/BB,EE}$ (e.g., [83,84], or Appendix B of G20), which take into account the effects of masking; the factor $f_{\ell} = [(\ell + 2)(\ell - 1)]/[\ell(\ell + 1)]$, which accounts for the mode-coupling matrix being applied to the shear field rather than to the convergence field directly, the pixel window function $F_{\ell}$, the nonlinear power spectrum $P_{NL}(\ell / \chi, \chi)$, modeled using HALOFIT as detailed in [85], and the reduced
skewness parameter $S_3$. The full derivation of $S_3$ is provided in Appendix A of G20, where it is evaluated to leading order in perturbation theory with the addition of a small-scale refinement (in the form of analytical fitting formulas) based on N-body simulations from [81]. In G20 we determined the range (i.e., angular scales and redshift interval) of validity of our $S_3$ model to ensure that modeling uncertainties will not affect our cosmological analysis. Since G20, however, some of our analysis choices changed; in particular, we updated the redshift distributions, the catalog shape noise, the measurement covariance, and the nuisance parameters priors, to reflect the updates in the DES Y3 data and modeling. Moreover, we include galaxy-galaxy lensing information from small scales in the form of shear ratios. Therefore, we repeated the modeling validation performed in G20 in Appendix B, using our updated analysis choices. Moreover, we validated our modeling on two different sets of N-body simulations (T17 and PKDGRAV).

A. Systematic effects

We model astrophysical and measurement systematic effects through nuisance parameters, over which we marginalize when estimating the cosmological parameters. Here is a short description of the nuisance parameters used in this work; priors are summarized in Table I.

Photometric redshift uncertainties. The first type of nuisance parameters are “calibration” parameters that model uncertainties in the photometric redshift estimates from the SOMPZ method. Such uncertainties are parameterized through a shift $\Delta z$ in the mean of the redshift distributions

$$n'(z) = \hat{n}'(z + \Delta z),$$

where $\hat{n}'$ is the original estimate of the redshift distribution for bin $i$. We assume DES Y3 priors for the shift parameters.

The priors also include the additional phot-z uncertainty due to blending [63]. This parametrization of the redshift uncertainties was shown to be adequate for the DES Y3 2-point analysis [15,86]; we nonetheless explore in Sec. VI a more complex parametrization of redshift uncertainties that also accounts for uncertainties in the shape of the redshift distributions.

Multiplicative shear biases. Biases coming from the shear measurement pipeline are modeled through an average multiplicative parameter $1 + m'_i$ for each tomographic bin. The effect of multiplicative shear biases on the measured moments can be modeled via,

$$\langle \kappa_{0,i}^{2} \rangle^{i,j} \rightarrow (1 + m'_i)(1 + m'_j) \langle \kappa_{0,i}^{2} \rangle^{i,j},$$

$$\langle \kappa_{0,i}^{3} \rangle^{i,j,k} \rightarrow (1 + m'_i)(1 + m'_j)(1 + m'_k) \langle \kappa_{0,i}^{3} \rangle^{i,j,k}. $$

We assume Gaussian priors for each of the $m'_i$ estimated following [63].

| Parameter | Prior |
|-----------|-------|
| $\Omega_m$ | U[0.1, 0.9] |
| $\sigma_8$ | U[0.5, 1.4] |
| $\Omega_b$ | U[0.03, 0.07] |
| $n_s$ | U[0.87, 1.07] |
| $h$ | U[0.55, 0.91] |
| $m_1$ | $\mathcal{N}(-0.0063, 0.0091)$ |
| $m_2$ | $\mathcal{N}(-0.0198, 0.0078)$ |
| $m_3$ | $\mathcal{N}(-0.0241, 0.0076)$ |
| $m_4$ | $\mathcal{N}(-0.0369, 0.0076)$ |
| $\Delta z_1$ | $\mathcal{N}(0.0, 0.018)$ |
| $\Delta z_2$ | $\mathcal{N}(0.0, 0.015)$ |
| $\Delta z_3$ | $\mathcal{N}(0.0, 0.011)$ |
| $\Delta z_4$ | $\mathcal{N}(0.0, 0.017)$ |
| $A_{IA,0}$ | U[-5, 5] |
| $a_{IA}$ | U[-5, 5] |
| $\Delta \varepsilon_{IA}$ | $\mathcal{N}(-0.009, 0.007)$ |
| $\Delta \varepsilon_{ZIA}$ | $\mathcal{N}(-0.035, 0.011)$ |
| $\Delta \varepsilon_{3}$ | $\mathcal{N}(-0.005, 0.006)$ |
| $\delta_{\varepsilon_{IA}}$ | $\mathcal{N}(0.975, 0.062)$ |
| $\delta_{\varepsilon_{ZIA}}$ | $\mathcal{N}(1.306, 0.093)$ |
| $\delta_{\varepsilon_{3}}$ | $\mathcal{N}(0.870, 0.054)$ |
| $b_{\gamma}$ | U[0, 3] |
| $b_{\beta}$ | U[0, 3] |
| $b_{\alpha}$ | U[0, 3] |

Intrinsic galaxy alignments (IA). We model IA following the nonlinear alignment (NLA) model [87–89]. It can be included in our modeling introducing $\delta_i = A(z)\delta$, which is the density contrast responsible for the intrinsic alignment, related to the matter density contrast $\delta$. In the NLA model, the IA amplitude can be written as a power law,

$$A(z) = -A_{IA,0} \left( \frac{1 + z}{1 + z_0} \right) a_{IA} c_{1P,0} \frac{D(z)}{D(z)},$$

where $D(z)$ is the angular distance to redshift $z$.
with \( z_0 = 0.62, c_1 \rho_{\text{crit}} = 0.0134 \), with \( \rho_{\text{crit}} = \rho_{m,0}/\Omega_m \) [87] and \( D(z) \) the linear growth factor [90]. For second moments, the NLA model can be incorporated in our theoretical predictions by modifying the lensing kernel,

\[
q^i (\chi) \rightarrow q^i (\chi) + A(z(\chi)) \frac{n^i(z(\chi))}{n^i} dz \frac{d\chi}{d\chi}. \tag{10}
\]

For third moments, we make the assumption that the NLA contribution follows the perturbation theory relation for the actual signal. [89] have shown this is in reasonable agreement with measurements from hydrodynamical simulations, so we follow them and modify Eq. (3) as follows:

\[
q^i q^j q^k \rightarrow q^i q^j q^k + \frac{A^2 + 2A}{3} (q^i q^j n^k + \text{cycl.}) + \frac{A^2 + 2A^3}{3} (q^i q^j n^k + \text{cycl.}) + A^4 (n^i n^j n^k), \tag{11}
\]

where in the above equation we dropped the redshift dependence for sake of simplicity; moreover, we used \( n = n^i(z(\chi)) \frac{dz}{d\chi} \), and cycl. refers to the cyclic permutation of the indexes \( i, j, k \) in the terms in parenthesis. We marginalize over \( A_{\text{IA}} \) and \( a_{\text{IA}} \) assuming flat priors. The fiducial DES Y3 \( 3 \times 2p \) analysis adopted a different, more general model for the intrinsic galaxy alignment, called “TATT” (Tidal Alignment and Tidal Torquing; [91]), that can capture the “tidal torquing” relevant for determining the angular momentum of spiral galaxies. Tidal torquing is ignored in the NLA model, which can account only for the tidal alignment of galaxies. We did not implement such a general model here; the DES Y3 cosmic shear analysis [18] found a weak preference for simpler IA modeling (i.e., for NLA rather than TATT), obtaining consistent cosmological constraints when different IA prescriptions were assumed. For this reason we use the NLA model as our fiducial choice.

Shear ratio parameters. We include in the analysis galaxy-galaxy lensing small scale information in the form of ratios of galaxy-galaxy lensing measurements [67]. These measurements use lenses as the first three tomographic redshift bins of the MAGLIM lens galaxy sample [92]. When modeling the shear ratio measurements, we marginalize over the uncertainties in the photo-z estimates of the lens samples through a shift \( \Delta z^{\text{lens}} \) in the mean of the redshift distributions and a stretch \( \delta^{\text{lens}} \) in their widths,

\[
n_{\text{lens},i}^{\text{lens},j} (z) = \delta^{\text{lens}} n_{\text{lens},i}^{\text{lens},j} (\delta^{\text{lens}} \{ z - (z) \} + \Delta z^{\text{lens}}), \tag{12}
\]

where \( (z) \) is the mean redshift of the lens sample. Priors on \( \Delta z^{\text{lens}} \) and \( \delta^{\text{lens}} \) are provided in [93]. We also marginalize over the galaxy-matter bias \( b^i_g \) of the three lens samples using broad flat priors.

B. Map making and moments estimator

We describe here how we measure the second and third moments of the convergence field starting from a weak lensing catalog. The following applies to both data and simulated catalogs, as they come in the same format.

Starting from the catalog, we first generate convergence maps for each tomographic bin. The convergence maps used in this work are estimated using a full-sky generalization of the [94] algorithm, first developed by [95]. The map-making process for the DES Y3 convergence maps is explained in full detail in [4], together with a thorough validation of the maps. Here, we briefly summarize the procedure.

We use the weak lensing catalog shear estimates to create pixelized maps for the two components of the shear field. The maps are constructed using healpix with \( \text{NSIDE} = 1024 \) (corresponding to a pixel size of 3.44 arcmin). The estimated value of the complex shear per pixel is given by

\[
\gamma_{\nu}^{\ell} = \sum_{ij=1}^{n} \epsilon_{ij} \omega_{ij}, \quad \nu = 1, 2, \tag{13}
\]

where \( \epsilon_{ij} \) is the per-galaxy observed ellipticity, \( \nu \) refers to the two shear field components, \( n \) is the total number of galaxies in the pixel, \( R \) is the average METACALIBRATION response of the sample (\( R = 1 \) for simulated catalogs), and \( \omega_{ij} \) is the per-galaxy inverse variance weight. The sum runs over all the galaxies in the pixel. Shear maps for each tomographic bin are created. As specified in Sec. III A, we do not explicitly correct for the multiplicative shear bias when making the maps, but rather we account for it during the cosmological inference. Any nonzero mean shear is subtracted from the catalog before creating the maps.

We then convert the shear maps into a curl-free E-mode convergence map \( \kappa_E \) and a divergence-free B-mode convergence map \( \kappa_B \) using a spin transformation. This is achieved by using the healpix function MAP2ALM to decompose the shear maps in spherical harmonic space obtaining the coefficients \( \hat{\gamma}_{E,\ell,m}, \hat{\gamma}_{B,\ell,m} \), and then calculating \( \kappa_{E,\ell,m}, \kappa_{B,\ell,m} \) as

\[
\gamma_{\ell,m} = \hat{\gamma}_{E,\ell,m} + i \gamma_{B,\ell,m} = -\sqrt{\frac{(\ell+2)(\ell-1)}{\ell+1}}(\kappa_{E,\ell,m} + i \kappa_{B,\ell,m}). \tag{14}
\]

Next we use the healpix function ALM2MAP to convert these coefficients back to real space \( \kappa_E \) and \( \kappa_B \) maps. The maps are smoothed using a top-hat filter and different smoothing scales \( \theta_0 \). In practice, this is achieved by multiplying the coefficients of the harmonic decompositions of the \( \kappa_E \) and \( \kappa_B \) maps by Eq. (5), prior to the conversion to real space. Simple estimators then give the moments of a smoothed map,
\[
\langle \hat{\kappa}^2_{\theta_0} \rangle_{i,j} = \frac{1}{N_{\text{tot}}} \sum_{\text{pix}} \hat{\kappa}_{\theta_0,\text{pix}}^2 \hat{\kappa}_{\theta_0,\text{pix}}, \tag{15}
\]
\[
\langle \hat{\kappa}^3_{\theta_0} \rangle_{i,j,k} = \frac{1}{N_{\text{tot}}} \sum_{\text{pix}} \hat{\kappa}_{\theta_0,\text{pix}}^3 \hat{\kappa}_{\theta_0,\text{pix}}^3 \hat{\kappa}_{\theta_0,\text{pix}}, \tag{16}
\]
where \(i, j, k\) refers to different tomographic bins. We estimate the moments for both the E- and B-mode convergence maps, although only the E-modes moments are used for the cosmological analysis. The sum runs over all the pixels on the sky (thus including regions outside the footprint). This is needed for two reasons: first, the transformation from the shear field to the convergence field is nonlocal and some power is transferred outside the footprint during the transformation, and second, the smoothing of the maps also transfers some of the power from the pixels close to the edge to pixels outside the footprint. We have shown in G20 that our modeling, together with the use of mode-coupling matrices, is able to take into account these effects (also including the lack of shear data outside the footprint, since the shear field is not defined there).

Due to the presence of shape noise, the measurement of galaxy shapes is only a noisy estimate of the shear field \(\gamma\). This also means that our estimate of the convergence field is noisy,

\[
\kappa_{E,\text{obs}} = \kappa_{E,\text{true}} + \kappa_{E,\text{noise}}, \tag{17}
\]
\[
\kappa_{B,\text{obs}} = \kappa_{B,\text{true}} + \kappa_{B,\text{noise}}. \tag{18}
\]
In the above equations, we omitted the smoothing angle \(\theta_0\). The contribution of the noise to the convergence field can be estimated by randomly rotating the shapes of the galaxies and applying the full-sky spherical harmonics approach to obtain the convergence [2,7]. As the random rotation should completely erase the cosmological contribution, the resulting convergence signal just contains noise and averages to zero (but with a non-negligible variance).

It follows that when estimating second and third moments from noisy convergence maps, it is necessary to properly denoise the measured moments. Following [7],

\[
\langle \kappa^2 \rangle_{i,j} \rightarrow \langle \kappa^2 \rangle_{i,j} - \langle \kappa \rangle_{\text{rand}} \langle \kappa \rangle_{i,j} - \langle \kappa_{\text{rand}} \rangle_{i,j} - \langle \kappa^2 \rangle_{i,j}, \tag{19}
\]
\[
\langle \kappa^3 \rangle_{i,j,k} \rightarrow \langle \kappa^3 \rangle_{i,j,k} - \langle \kappa^3 \rangle_{i,j,k} - \langle \kappa_{\text{rand}} \rangle_{i,j,k} - \langle \kappa^2 \rangle_{i,j,k} + \text{cycl}, \tag{20}
\]
where cycl. refers to the cyclic permutation of the indexes \(i, j, k\) for the terms in parenthesis. In the above equations, the term \(\langle \kappa^2 \rangle_{i,j} \langle \kappa_{\text{rand}} \rangle_{i,j} \langle \kappa^2 \rangle_{i,j} \) is the noise-only contribution to the second (third) moments of the tomographic bins \(i, j, (k)\). Under certain conditions, most of these terms vanish; those terms that do not vanish need to be subtracted from the measured moments. We verified which terms vanish in Appendix D.

**IV. LIKELIHOOD AND COVARIANCE**

This section provides details about our data vector, likelihood, and covariance. Our data vector consists of all the possible combinations of second and third moments involving the four weak lensing tomographic bins. This adds up to a total of ten combinations of second moments and 20 combinations of third moments. For each of these second and third moments, we consider 10 equally (logarithmic) spaced smoothing scales \(\theta_0 \in [3.2, 200]\) arcmin. We then remove scales following G20, i.e., we remove angular scales smaller than a corresponding comoving scale \(R_0\) given by \(\theta_0 = R_0/\chi(\bar{z})\), where \(\chi(\bar{z})\) is the comoving distance of the mean redshift of a given tomographic bin. In the case of moments from different tomographic bins, we took the average of the mean \(\bar{z}\) of the two bins. This scale cut is designed to remove scales significantly affected by modeling uncertainties that could contaminate the cosmological analysis, with the dominant uncertainty being contamination due to baryonic effects. G20 determined the fiducial scale cut to be 24\(h^{-1}\) Mpc when combining second and third moments. We adopt here a scale cut of 28\(h^{-1}\) Mpc. This change is necessary because the simulated analysis in G20 did not use the final setup for the analysis (e.g., inclusion of the shear-ratio likelihood, final values for redshift distributions, shape noise, effective number densities, covariance, etc.); we therefore repeated the scale cut analysis with all the analysis ingredients updated, and determined 28\(h^{-1}\) Mpc to be the correct scale cut to be used in this analysis (see Appendix A for more details).

We then compress our data vector using the massively optimized parameter estimation and data compression (MOPED) algorithm [96–98] based on the Karhunen-Loève algorithm, which allows us to reduce the dimensionality of our data vector to the number of model parameters considered. In our case, the number of parameters used to model the moments data vector is 15; therefore, the size of the compressed moments data vector is 15. The compression allows us to reduce the enlargement of the parameters posterior due to noise in the precision matrix estimate, as the covariance matrix is estimated from a limited number of simulations [99]. The final enlargement depends on the size of the compressed data vector rather than on the size of the uncompressed data vector, which makes having an efficient compression scheme desirable. In particular

\[
d_i^{\text{compr}} = \langle d_i \rangle_C^{-1} C^{-1} d \equiv b_i d, \tag{21}
\]
where \(d\) is the full-length data vector, \(\hat{C}\) is the measurement covariance, and \(d_i^{\text{compr}}\) is the \(i\)-th element of the compressed data vector. The index \(i\) refers to the \(i\)-th model parameter \(p\)
considered, and \((df)\) is the derivative of the model data vector with respect to that parameter.

We evaluate the posterior of the parameters conditional on the data by assuming a Gaussian likelihood for the data, i.e.,

\[
-2\ln L = f_2f_1[\hat{d} - M(p)]\hat{C}^{-1}[\hat{d} - M(p)]^T. \tag{22}
\]

Here \(M(p)\) is our theoretical model, \(\hat{d}\) is the data vector, and \(\hat{C}^{-1}\) is the inverse of our covariance estimate. The posterior is then the product of the likelihood and the priors. Note that the quantities \(M(p), \hat{d}\) and \(\hat{C}^{-1}\) in Eq. (22) are to be considered compressed quantities. The terms \(f_1\) and \(f_2\) account for noise introduced when the covariance matrix is estimated from random realizations of the data \([99–101]\) and are given by

\[
f_1 = \frac{N_{\text{sims}} - N_{\text{data}} - 2}{N_{\text{sims}} - 1}, \tag{23}
\]

\[
f_2 = \left[ 1 + \frac{(N_{\text{data}} - N_{\text{pars}})(N_{\text{sims}} - N_{\text{data}} - 2)}{(N_{\text{sims}} - N_{\text{data}} - 1)(N_{\text{sims}} - N_{\text{data}} - 4)} \right]^{-1}, \tag{24}
\]

where in our case the number of independent realizations used to estimate the covariance is \(N_{\text{sims}}\) (i.e., the number of independent simulations) and \(N_{\text{data}}\) is the length of the data vector. In the case of compressed quantities, \(f_1, f_2 \sim 1\) as \(N_{\text{sims}} \gg N_{\text{data}}\).

To correctly infer cosmological parameters from our data, we need an accurate estimate of the measurement uncertainty. Our fiducial method to estimate the covariance uses 1000 independent realizations of the convergence maps generated from the FLASK simulations. As an additional check, we also estimate the covariance using the PKDGRAV and T17 simulations. The PKDGRAV and T17 simulations (Fig. 2) have been produced at cosmologies different to that of the FLASK simulations; hence, these alternative covariances provide extra validation against the dependency of our covariance on the value of cosmological parameters. More details are given in Appendix G. Given a set of N-body simulations, for each realization we measure the second and third moments of the smoothed convergence field and build the covariance matrix as

\[
\hat{C} = \frac{1}{\nu} \sum_{i=1}^{N_r} (\hat{d}_i - \bar{d})(\hat{d}_i - \bar{d})^T, \tag{25}
\]

where \(\nu = N_r - 1\) with \(N_r\) the number of realizations, \(\hat{d}_i\) the data vector measured in the \(i\)th simulation, and \(\bar{d}\) the sample mean. The data vector is made of a combination of second and third moments as measured at different smoothing scales. We also add to our covariance a “modeling uncertainty” related to the analytical fitting formulas describing the third moments at small scales (see G20 for more details). We then compress the covariance following

![FIG. 2. Top: measured correlation matrix of second and third moments from FLASK simulations (lower right triangle) and from T17 simulations (upper left triangle). No scale cut has been applied. From bottom left to top right, we show the bins: \([1, 1], [2, 2], [3, 3], [4, 4], [1, 2], [1, 3], [1, 4], [2, 3], [2, 4], [3, 4], [1, 1], [1, 2], [2, 2], [3, 3], [4, 4], [1, 1], [2, 1], [3, 1], [3, 2], [4, 1], [2, 3], [2, 4], [3, 4], [4, 4], [1, 2], [1, 3], [1, 4], [2, 3], [2, 4], [3, 4], [4, 4], [1, 2], [1, 3], [1, 4], [2, 3], [2, 4], [3, 4], [4, 4], [1, 2], [1, 3], [1, 4], [2, 3]). The main difference between the two covariances is that the FLASK one has on average 5–10% larger amplitude at large scales. The difference vanishes at small scales since those are dominated by shape noise. Bottom: same as the top image, but focusing on a few elements (from bottom left to top right, we show \([1, 1], [1, 4], [4, 4]\) for second moments and \([1, 1], [1, 4], [4, 4]\) for third moments), and showing only the dynamical range \([-0.1, 1.0]\) effectively spanned by the elements of the correlation matrix. The diagonal blocks of the correlation matrix contain essentially all the non-negligible contributions. On large scales, where the cosmic variance contribution to the covariance dominates, the window function of the moments generates off-diagonal terms (within the block diagonal matrix) while on small scales these are due to the pixel window function. Note that in the absence of shape noise we also expect a contribution from nonlinear evolution on small scales. The third moments correlation matrices are more diagonal than second moments ones owing to a larger shape noise contribution. Non-negligible cross-covariance between second and third moments is limited to very large scales, and is generally small (10–15% at most).}
\[ C_{ij}^{\text{compr}} = b_i^T \hat{C} b_j. \]  

We tested that using the FLASK covariance we were able to correctly recover the input cosmology in simulations (Appendix B). In the inference, we also add an independent “shear ratio” likelihood [67]. The shear ratio likelihood uses small scale information from the ratio of galaxy-galaxy lensing measurements (the mean tangential shear around lens galaxies) between two weak lensing source tomographic bins and a shared lens sample. Its inclusion improves the constraints on the redshift distributions and on other nuisance parameters of our model. The shear ratio data vector consists of nine scale-averaged ratios. We use as a lens the first three tomographic redshift bins of the MAgLIM lens galaxy sample [92]. The shear ratios likelihood is modeled as an independent Gaussian likelihood, and uses an analytical covariance matrix. The assumption of independence is justified by the smallness of the scales involved in the shear ratio measurements (less than 6 h^{-1} Mpc). We note that the scale cut for this work is 28 h^{-1} Mpc, although the two scale cut limits cannot be directly compared since the mass map smoothing function and the galaxy-galaxy lensing angular bin kernels weight scales slightly differently. Nonetheless, the independency of the shear ratio likelihood has been proven in the context of the DES Y3 3 × 2pt analysis [67]. Because we adopt the same scale cut criteria as the DES 3 × 2pt analysis, we assume independency holds here as well. Lastly, since the shear ratio covariance is analytical, we do not compress the shear ratio data vector.

Having defined the likelihood, we sample the posteriors of our parameters using POLYCHORD [102,103]; this is a nested sampler that uses slice sampling within the nested isolikelihood contours. For the cosmological parameters, we assume a flat $\Lambda$CDM cosmology and vary five parameters; $\Omega_m$ (the density of the total matter today), $\sigma_8$ (the amplitude of structure fluctuations in the present day Universe, parameterized as the standard deviation of the linear overdensity fluctuations on a 8 h^{-1} Mpc scale), $\Omega_b$ (the baryonic density in units of the critical density), $n_s$ (the spectral index of primordial density fluctuations), and $h$ (the dimensionless Hubble parameter). We assume wide flat priors on $\Omega_m$ and $\sigma_8$ and adopt the informative priors on $h$, $n_s$, and $\Omega_b$ that were used in the DES Y3 2-point function 3 × 2pt analysis (see Table I). When constraining cosmological parameters, we marginalize over nuisance parameters describing mean photo-$\pi$ uncertainties, multiplicative shear biases and IA effects in our measurements. The modeling of our nuisance parameters is described in Sec. III A. Photo-$\pi$ uncertainties are parametrized by a shift in the mean of the distribution (one for each tomographic bin). Priors for the shifts come from Myles and Alarcon et al. [65]. Multiplicative shear bias priors are described in MacCrann et al. [63]. We also assume wide flat priors for intrinsic alignment amplitudes. The addition of the shear-ratio likelihood to the analysis necessitates additional modeling parameters, summarized in Table I. These are lens redshift parameters (modifying the mean redshift and the width of the lens sample redshift distributions) and one free (linear) galaxy bias parameter per lens bin.

Last, we note that since the theory predictions described in Sec. III are time consuming to compute due to the large number of cross-correlations and integrations involved, we implemented an emulator [104,105] to speed up the calculations. In our implementation, the emulator provides fast theoretical predictions by interpolating over a number of predictions computed at a set of training points spanning the parameter space of interest (in our case, the five cosmological parameters). In particular, the quantities emulated are the terms

\[ \langle \delta_{h_0}^2 \rangle_{\text{EE/BB}} (\chi) \equiv \sum_{\ell} \frac{2\ell + 1}{4\pi} f_{\ell} W_\chi (\theta_0)^2 \times \sum_{\ell''} M_{\ell \ell''}^{\text{EE/BB}} P_{\delta \chi} (\ell', \ell'') \chi f_{\ell'}, \]  

which enter in the modeling of Eqs. (2) and (3). The accuracy of the emulator is sufficient to not bias the cosmological analysis, as demonstrated in G20.

V. PREUNBLINDING TESTS

Before proceeding to unblind the data vector and analyse the results of the unblinded analysis, we performed a number of tests. These tests complement the ones performed at the catalog and map level presented in [4,14]. We remind the reader that when this analysis was performed, the shape catalog was already deemed science ready and unblinded, and only the data vector level of blinding was enforced. The whole cosmological pipeline had already been demonstrated in G20 to recover the true cosmology using realistic simulations. We nonetheless repeated the validation in simulations with the updated analysis choices (e.g., redshift distributions, shape noise, priors, etc.) in Appendix B, using both T17 and PKDGRAV simulations. We also slightly changed the scale cut decided in G20, due to updates in the analysis choices. More details concerning the scale cuts are given in Appendix A.

We first performed two tests at the data vector level:

(i) We checked that additive biases due to PSF modeling errors were negligible at the data vector level, i.e., if neglected they would not bias our cosmological analysis. This test is similar to the test performed for the DES Y3 cosmic shear analysis [15]; more details are given in Appendix C.
(ii) We tested that mixed moments of convergence maps E-mode and noise (e.g., (κ_ν)_{XY}) are consistent with expectations based on tests on N-body simulations; more details are given in Appendix D. We then ran our analysis on blinded data vectors, and checked that:

(i) Cosmological constraints obtained using (blinded) second and third moments were consistent with each other. To this aim, we used posterior predictive distributions (PPD, [106]); see Appendix F.

(ii) The (blinded) posteriors of the systematic parameters did not concentrate at the edge of the prior. The level of agreement/disagreement with the prior was tested using a Gaussian estimator called the “update difference-in-mean” (UDM) statistic [107] (Appendix H).

We then unblinded the data vectors and ran the fiducial analysis; before looking at the unblinded posteriors, we further checked that:

(i) The goodness-of-fit p—value on unblinded data vectors was larger than 1%; see Sec. VI.

(ii) The best-fitting cosmology provided a good description to second and third moments B-modes (which are not included in the data vector), see Appendix E. This was done in an automated fashion such that we did not look at the actual best-fitting values.

In order to quantify goodness-of-fit and internal consistency among different parts of our data vector, we use the PPD methodology developed by [106] and adopted in the main DES Y3 3 × 2pt analysis. The PPD methodology derives a calibrated probability-to-exceed p; in the case of goodness-of-fit tests, this is achieved by drawing realizations of the data vector for parameters drawn from the posterior under study; for consistency tests (e.g., second moments vs. third moments), the realizations are drawn from disjoint subsets of the data vector. These realizations are then compared to actual observations and a distance metric (χ^2) is computed in data space, which is then used to compute the p—value. Once all these tests were passed, we looked at the unblinded posteriors of our analysis.

VI. COSMOLOGICAL CONSTRAINTS

We present here the cosmological constraints obtained assuming the ΛCDM model, varying five cosmological parameters and 19 nuisance parameters (10 for the moments likelihood and 9 additional ones for the shear ratio likelihood), as summarized in Table I. In addition to these parameters, we will also quote results in terms of the S_8 parameter, defined as

\[ S_8 \equiv \sigma_8 (\Omega_m / 0.3) \alpha. \]  

The value of \( \alpha \) can be chosen such that \( S_8 \) best constrains the degeneracy between \( \Omega_m \) and \( \sigma_8 \). However, the second and third moments have a slightly different degeneracy direction and so there is no value of \( \alpha \) that simultaneously optimizes both. For sake of simplicity we adopt \( \alpha = 0.5 \).

Figure 4 shows the posteriors for \( S_8, \Omega_m \), and \( \sigma_8 \) from the second and third moments individually, and from the combinations of the two. Third moments are much less constraining than second moments alone, but they are characterized by a slightly different degeneracy tilt in the \( \sigma_8-\Omega_m \) plane compared to second moments. The marginalized mean values of \( S_8, \Omega_m \), and \( \sigma_8 \) for the combination of second and third moments, along with the 68% confidence intervals, are

\[ \Omega_m = 0.27 \pm 0.03, \]  
\[ \sigma_8 = 0.83 \pm 0.05, \]  
\[ S_8 = 0.784 \pm 0.013. \]  

We report the constraints from the analysis of second and third moments individually in Table II, and for \( S_8 \) we additionally provide a visual comparison in Fig. 5. The combined moments analysis places a 1.7% constraint on \( S_8 \) and a 10% constraint on \( \Omega_m \), improving by ~15% and ~25% over constraints from second moments only. This level of improvement is expected (G20), and is due to the additional non-Gaussian information probed by third moments and the degeneracy breaking when second and third moments are combined. Table II also reports the p-values for the goodness-of-fit tests; these are well above the p-value = 0.01 threshold. The unblinded data vectors, along with the best-fitting models from our posteriors, are shown in Fig. 3. We caution the reader from any \( \chi^2 \)-by-eye estimate, as the different scales are highly correlated (especially for second moments, where adjacent scales have a correlations higher than 90%). Constraints from second and third moments are consistent with each other, although it is evident from Fig. 4 that they probe different parts of the parameter space in the \( \sigma_8-\Omega_m \) plane.

In Appendix F we use PPD to quantify the internal consistency of our data sets. In particular, we tested the compatibility between second and third moments constraints, between small and large scales, and between parts of the data vector using different redshift bins. These tests were performed prior to unblinding, using blinded data vectors, and were repeated after unblinding (although only the compatibility of second and third moments was considered as an unblinding criterion). In Appendix F we also perform a test analysing the data vector using a different parametrization of the redshift uncertainties, called hyperrank [86].

The results reported here were obtained using the FLASK covariance; in addition, we tested that our results do not change significantly when using the covariances estimated using the T17 or PKDGRAV simulations (Appendix G).
TABLE II. Constraints on the cosmological parameters $S_8$, $\Omega_m$, and $\sigma_8$. For each parameter we report the mean of the posterior and the 68% confidence interval. For the fiducial results (second moments, third moments, and the combination of the two) we also report the PPD goodness-of-fit $p$-value.

|                  | $S_8$        | $\Omega_m$   | $\sigma_8$ | $p$-value |
|------------------|-------------|--------------|------------|-----------|
| **Fiducial**     |             |              |            |           |
| Second moments   | 0.799 ± 0.015 | 0.21 ± 0.04  | 0.98 ± 0.10 | 0.21      |
| Third moments    | 0.72 ± 0.05  | 0.33 ± 0.16  | 0.73 ± 0.16 | 0.63      |
| Second + third   | 0.784 ± 0.013 | 0.27 ± 0.03  | 0.83 ± 0.05 | 0.26      |
| moments, no bin 1| 0.785 ± 0.014 | 0.30 ± 0.04  | 0.79 ± 0.06 |           |
| Second + third   | 0.779 ± 0.015 | 0.27 ± 0.04  | 0.83 ± 0.06 |           |
| moments, no bin 2| 0.789 ± 0.019 | 0.27 ± 0.05  | 0.83 ± 0.08 |           |
| Second + third   | 0.791 ± 0.018 | 0.23 ± 0.04  | 0.92 ± 0.08 |           |
| moments, no bin 3| 0.799 ± 0.014 | 0.26 ± 0.03  | 0.83 ± 0.05 |           |
| Second + third   | 0.780 ± 0.017 | 0.32 ± 0.05  | 0.76 ± 0.07 |           |
| moments, small   | 0.76 ± 0.02  | 0.28 ± 0.04  | 0.79 ± 0.07 |           |
| scales           |             |              |            |           |
| Second + third   | 0.782 ± 0.017 | 0.27 ± 0.04  | 0.83 ± 0.06 |           |
| moments, large   | 0.785 ± 0.015 | 0.27 ± 0.03  | 0.82 ± 0.06 |           |
| scales           | 0.788 ± 0.015 | 0.28 ± 0.03  | 0.82 ± 0.06 |           |
| **Variations**   |             |              |            |           |
| DES Y3 Cosmic    | 0.772 ± 0.016 | 0.29 ± 0.05  | 0.79 ± 0.08 |           |
| Shear, TATT free | 0.788 ± 0.016 | 0.28 ± 0.04  | 0.82 ± 0.08 |           |
| neutrino [15,18] | 0.779 ± 0.014 | 0.33 ± 0.03  | 0.74 ± 0.04 |           |
| DES Y3 Cosmic    | 0.797 ± 0.014 | 0.28 ± 0.07  | 0.85 ± 0.11 |           |
| Shear, NLA fixed | 0.751 ± 0.021 | 0.29 ± 0.08  | 0.79 ± 0.13 |           |
| neutrino [108]   |             |              |            |           |
| DES Y3 3 × 2pt, | 0.778 ± 0.031 | 0.18 ± 0.07  | 1.05 ± 0.16 |           |
| TATT free neutrino [108] |       |           |            |           |
| Other works      | 0.834 ± 0.016 | 0.316 ± 0.008 | 0.812 ± 0.007 | |  

A. Intrinsic alignment constraints and impact of the shear ratio likelihood

Intrinsic alignment (IA) is a potentially important contribution to the shear signal. We show in Fig. 6 the posterior of the IA amplitude parameter $A_{IA}$ for the combination of second and third moments. Our results are compatible with a null IA signal, as the amplitude of the IA signal is constrained to $A_{IA} = −0.09 ± 0.17$. Most of the constraint on IA comes from the shear ratio likelihood (Fig. 6), although when performing the analysis without shear ratio we also obtain a null IA signal of $A_{IA} = 0.09 ± 0.6$. The improvement in the IA constraints due to the inclusion of shear ratio is expected [67]; moreover, because of the slight degeneracy between the IA amplitude parameter and $S_8$, shear ratio also improves the $S_8$ constraints (~25%). The constraints obtained analysing second and third moments only are also very similar: $−0.08 ± 0.17$ and $−0.10 ± 0.15$ for second and third moments, respectively. The tighter constraint on $A_{IA}$ from third moments is due to a projection effect related to the broader constraints on $\Omega_m$. These results are compatible with the DES Y3 cosmic shear and $3 \times 2pt$ analyses results [15,18,108], which also find an IA amplitude consistent with zero. Lastly, we ran an additional test analysing our data vector assuming no IA ($A_{IA} = 0$); the results are shown in Fig. 6 and are almost identical to the fiducial results. The only difference between the no IA model and the fiducial analysis is that the former strongly constrains the nuisance parameter $\Delta z_1$ (redshift uncertainty of the first redshift bin). In particular, the posterior of that parameter is shrunk by half, although it is still consistent with zero. The IA model (NLA) used in this work is simpler than the fiducial model (TATT) adopted by the DES Y3 $3 \times 2pt$ analysis [91]. However, [18] finds that simpler IA models such as NLA are sufficient for modeling the DES Y3 data, so we do not think any of the conclusions in this work are affected by our (simpler) IA modeling choice.

B. Comparison with DES constraints

We discuss here how the parameter constraints obtained from this work compare with the ones obtained by other cosmological analyses using DES Y3 data (cosmic shear, $3 \times 2pt$, and lensing peaks). Marginalized posteriors for $S_8$ and $\Omega_m$ are shown in Fig. 7 and (for $S_8$ only) in Fig. 5; we also report the numerical values in Table II. While the level of agreement can be noted in these figures, we cannot quantify it using the PPD metric, as we do not have the cross-covariance of moments with the other data vectors (a requirement of the PPD method).

The comparison that is probably the most relevant is with cosmic shear, which is a two-point correlation of the same lensing field. Our second moment should be consistent with it, although as discussed below the weighting of different scales (in particular in Fourier space) differs. The peaks statistic uses different non-Gaussian information from the third moment, so that is an interesting comparison as well. For completeness we include the $3 \times 2pt$ results although these use the clustering of lens galaxies (a different probe of the mass distribution). However within the context of
FIG. 3. Measured second moments (upper plots) and third moments (lower plots). Red points represent the measurement. Gray shaded regions highlight the scales removed by the analysis. The conservative scale cut implemented in this analysis removes a large part of our data vector. Solid, dotted, and dashed lines represent the predictions obtained using the best-fitting cosmology of second and third moments analysis (either considered in combinations or alone). Data points are very correlated (Fig. 2), so we caution the reader from any $\chi^2$-by-eye estimation.
ΛCDM, the results should agree provided the theoretical predictions are accurate and the mitigation of systematic errors in each analysis is reliable.

**DES Y3 cosmic shear.** The first comparison with other DES Y3 constraints is with the cosmic shear analysis \[15,18\]. We compare with the constraints from two slightly different cosmic shear analyses: the first one is a ΛCDM analysis which assumes a more complex IA model (the TATT model), and marginalizes over the neutrino mass, whereas the second one, which better matches the analysis choices adopted in this work, assumes NLA as IA model and fixes the neutrino mass to zero.\(^1\) Both adopt the DES Y3 ΛCDM optimized scale cut.\(^2\) The constraints from the combination of second and third moments are in good agreement with the constraints from both cosmic shear analyses.

In terms of constraining power on the \(S_8\) and \(\Omega_m\) parameters, the NLA + fixed neutrino cosmic shear analysis is similar to the analysis using second moments only. The combined moments analysis is more constraining, due to the additional non-Gaussian information and the degeneracy breaking of the third moments. Although both cosmic shear and second moments are Gaussian statistics and they both probe the shear power spectrum, their posteriors do not have to perfectly overlap, as they weight power spectrum multipoles differently (Appendix I). In particular, our scale cuts exclude some of the higher wave number contributions to \(\xi_{++}\). Nonetheless, the peaks of the second moments and the combination of second and third moments posteriors are consistent with the peak of the NLA + fixed neutrino cosmic shear posterior in the \(S_8-\Omega_m\) plane (1σ and 0.15σ, respectively).

**DES Y3 3 × 2pt.** Similar to the DES Y3 cosmic shear analysis, we compare to two different versions of the DES Y3 3 × 2pt analysis \[108\]: a first one that assumes ΛCDM,
the TATT model and marginalizes over the neutrino mass, and a second one which better matches the analysis choices adopted in this work, assuming NLA as IA model and fixing the neutrino mass to zero. We report the latter analysis for a visual comparison of the constraining power, but we caution the reader that it is unlikely to pass our scale cuts criteria, which impose a maximum bias of 0.3 σ in the $S_8-\Omega_m$ plane in case of baryonic contamination. This analysis was not presented in [108], and no adequate scale cut was determined. For sake of simplicity, we decided to use the same scale cut adopted in the $3 \times 2pt$ TATT + free neutrino mass analysis, which is likely too aggressive. This is because we know that the $3 \times 2pt$ TATT + free neutrino mass analysis passes the scale cuts criteria with exactly a 0.3σ bias [108]; the NLA + fixed neutrino analysis, having slightly more constraining power, is likely to fail those criteria. To avoid misinterpreting these results, we decided to shift the contours to lie on top of the $3 \times 2pt$ TATT posterior, such that the real position is unknown and the posterior can only be used to get a sense of the effect of different analysis choices on the constraining power of the $3 \times 2pt$ analysis. The $3 \times 2pt$ analysis relies on three different probes: cosmic shear, galaxy-galaxy lensing, and galaxy clustering. Remarkably, the $S_8$ constraining power from the moments analysis is 10% better than that from the DES Y3 $3 \times 2pt$ analysis, despite not relying on a lens sample or the $2 \times 2pt$ part of the data vector. The DES Y3 $3 \times 2pt$ constraints are, however, slightly more stringent in terms of $\Omega_m$ (by 10%), due to the significant contribution from the galaxy-galaxy lensing and galaxy clustering part of the analysis. The posteriors show good overlap, with the moments peak being $\sim 1.1\sigma$ away from the DES Y3 $3 \times 2pt$ TATT + free neutrino analysis peak in the $S_8-\Omega_m$ plane. Given that the constraints come from different probes we can consider the posteriors to be in reasonably good agreement.

**DES Y3 Peaks + Power spectrum analysis.** [79] use peak counts to extract non-Gaussian information from the convergence field, and combine this with constraints from the power spectrum of convergence maps. The comparison of our analysis with theirs is interesting for two reasons: (1) similar to this analysis, it exploits some non-Gaussian information of the convergence field to constrain cosmological parameters; (2) the Peaks + Power spectrum

---

**FIG. 5.** Constraints on the cosmological parameter $S_8$; we report the mean of the posterior and the 68% confidence interval. The first three lines are the fiducial results from this analysis. Following that are the $S_8$ values for a number of variations and systematic tests; removing one redshift bin at a time, using a different parametrization for the redshift distribution called “hyperrank”, considering only small or large scales, considering the case with no intrinsic alignment or no shear ratio (SR), and using different covariance matrices (FLASK + T17 or FLASK + PKDGRAV), as explained in Appendices F and G. Lastly, we compare with a number of results from other works, either with DES or external data.
analysis uses an independent, completely different framework to provide theory predictions for the observables—they forward model the measurements using a Gaussian process emulator built using N-body simulations of different cosmologies. The analysis choices of the Peaks + Power spectrum analysis and our moments analysis are very similar, the main difference being that the former does not use the shear ratio likelihood and uses somewhat tighter priors for the $n_s$, $h_{100}$, and $\Omega_b$ parameters. The results from these two analyses are in agreement (Fig. 7), with the peaks of their posterior within 1σ of ours in the $S_8$-$\Omega_m$ plane.

C. Comparison with external data sets

We compare here our parameter constraints with the results obtained using external data sets. In particular, we compare with the recent results of the KIDS-1000 survey [21], HSC [16], and Planck [111]. In order to estimate the tension between different analyses, we calculate a Monte Carlo estimate of the probability of a parameter difference [107,112], using the TENSOMETER software. In the case of uncorrelated data sets, the probability of the parameter difference reads

$$P(\Delta \theta) = \int_{V_P} P_A(\theta) P_B(\theta - \Delta \theta) d\theta,$$

(33)

where $V_P$ is the prior support and $P_A$ and $P_B$ are the two posterior distributions of the parameters. The probability of an actual shift in parameter space is obtained from the density of parameter shifts,

$$\Delta = \int_{P(\Delta \theta) > P(0)} P(\Delta \theta) d\Delta \theta.$$

(34)
which is the posterior mass above the contour of constant probability for no shift, \( \Delta \theta = 0 \). Due to the discrete nature of our posterior samples, the integral in Eq. (34) is evaluated using a Monte Carlo approach [112].

A visual comparison between the results of the moments analysis and the results obtained from external data sets is provided in Fig. 8 for the \( S_8 \) and \( \Omega_m \) parameters and in Fig. 5 for \( S_8 \) only; additionally, the probability of the parameter difference is reported in Table III. The moments analysis is in good agreement with the other weak lensing analyses considered here (\( \lesssim 1\sigma \)), and it is the most constraining one (owing both to the larger data set and to the extra non-Gaussian information probed by the moments).

When comparing with the results from the Planck analysis, however, we measure a larger tension, at the level of \( 2.2 - 2.8\sigma \), depending on the combination of moments considered (see Fig. 8). The third moment independently is in \( 2.8\sigma \) tension with Planck, which provides a cross-check on the other analyses of 2-point correlations. Note that the joint constraint, though tighter, is in slightly lower tension. Interestingly, the moments analysis is significantly more constraining than Planck for the \( S_8 \) parameter.

When comparing results from different analyses, we did not try to unify different analysis choices (e.g., priors, scale cuts, etc.); this complicates the comparison [113]. Nevertheless, the moments analysis, in line with other weak lensing analyses, favours lower \( S_8 \) values than Planck.

VII. SUMMARY

We presented a cosmological analysis of the second and third moments of weak lensing mass (convergence) maps from the Y3 data of the DES. The second moment of the convergence as a function of smoothing scale contains information similar to standard shear 2-point statistics, whereas the third moment, or skewness, contains additional non-Gaussian information. Several theoretical studies have explored the use of statistics beyond 2-point correlations to extract additional non-Gaussian information from lensing data. The 3-point function is the lowest order statistic in perturbation theory and is the simplest to model and interpret. Its signal-to-noise is significantly smaller than for 2-point correlations, but its dependence on the key cosmological parameters \( (S_8, \Omega_m) \) differs, enabling partial degeneracy breaking and improved constraints on cosmological parameters. Our study is the first to test these theoretical expectations with data in a comprehensive way, following an end-to-end analysis of mock catalogs that included the expected leading sources of systematic uncertainty (see G20). We note that the counts of peaks in the lensing field are analyzed in a separate DES paper [79] and other non-Gaussian statistics such as the topological

TABLE III. Probability of the parameter difference (computed over the full parameter space) between the DES Y3 moments analysis and three analyses using external data sets: KIDS-1000 survey [21], HSC [16], and Planck [111].

| Planck TTTEEE Power spectrum | HSC Y1 lowl lowE | KIDS-1000 | 
|-------------------------------|------------------|------------|
| Second moments | 2.7\( \sigma \) | 0.3\( \sigma \) | 0.9\( \sigma \) |
| Third moments | 2.8\( \sigma \) | 1.2\( \sigma \) | 0.2\( \sigma \) |
| Second + third moments | 2.2\( \sigma \) | 0.9\( \sigma \) | 0.6\( \sigma \) |
Minkowski functionals as well as deep learning approaches have been proposed as well (see Sec. I for a review).

Our analysis relies on 100 million galaxy shapes measured over 4139 square degrees, which have been used to reconstruct the convergence field in four source redshift bins. The data has been analyzed in the context of the ΛCDM model, varying five cosmological parameters (Ω_m, σ_8, n_s, Ω_b, and h_{100}) and 19 nuisance parameters (modeling astrophysical and measurement uncertainties). One of our goals is to quantify the tension between CMB and late time estimates of S_8 and other relevant parameters. In view of several recent measurements reporting tension between the amplitude of mass fluctuations in the late times vs. early universe (as probed by the CMB), we have carried out measurements and consistency tests of ΛCDM rather than pursue extended cosmological models. The modeling used to describe the second and third moments measured in data is analytical: as described in G20 we have built an emulator to obtain rapid predictions from perturbation theory calculations well tested with N-body simulations. Thus the cosmological analysis here does not rely on large suites of N-body simulations to forward model the signal.

The combined analysis of second and third moments was able to constrain S_8 ∝ σ_8(Ω_m/0.3)^0.5 with 1.7% uncertainty and Ω_m with 10%: in particular, we obtained S_8 = 0.784 ± 0.013 and Ω_m = 0.27 ± 0.03. The third moments improved the constraints on S_8 and Ω_m by ∼15% and ∼25%, respectively, in line with the expectation based on simulations (G20). The improvement is due to the degeneracy breaking and the non-Gaussian information probed by the third moments. The goodness-of-fit p—value of the data vectors (second, third, and the combination of second and third moments) was found to be way larger than 1%, which is our criterion for a reasonable goodness-of-fit.

We performed our analysis following the blinding scheme proposed by [68]. Before unblinding the analysis, we performed a number of systematic tests which had been defined as unblinding criteria: we checked that additive biases due to PSF modeling errors were small enough to not bias the cosmological analysis; that mixed moments between convergence map E-modes and noise were consistent with expectations based on tests on N-body simulations; that cosmological constraints obtained using second and third moments were consistent with each other using posterior predictive distributions (PPD, [106]); that the best-fitting cosmology provided a good description of the B-modes of the second and third moments as well (the B-modes were not included in the data vector used for the cosmological analysis); that the posteriors of the nuisance parameters did not concentrate at the edge of the prior, tested using the Gaussian estimator update difference-in-mean (UDM) statistic [107]. All these tests were successfully passed. After unblinding, we further used PPD to assess the internal consistency of other subsets of the data vector (small vs. large scales, or across redshift bins); we also tested that our results were robust against different modeling choices for the covariance matrix used in the analysis, or the inclusion of small-scale galaxy-galaxy lensing ratios (a.k.a. shear ratios, [67]). All tests performed after unblinding validated the robustness of our results.

Constraints from the combination of second and third moments were found to be compatible with constraints from the DES Y3 cosmic shear analysis [15,18], the DES Y3 3 × 2pt analysis [108], and the DES Peaks + Power spectrum analysis [79]. In terms of constraining power, the addition of non-Gaussian information via the third moments in the analysis may be regarded as successful—the constraints on S_8 and Ω_m were shown to be tighter than from DES cosmic shear, and, for S_8, similar to the DES 3 × 2pt constraint.

We compared our constraints to two contemporaneous lensing surveys: the KIDS-1000 survey [21] and the HSC Y1 data [16], finding agreement (≤1σ).

When compared to predictions based on CMB data from the Planck satellite [111], we estimate a 2.2 ± 2.8σ tension in the full parameter space, depending on the combination of moments considered (see Table III). The moments analysis favours lower S_8 values compared to Planck, in line with other weak lensing analyses. Interestingly the third moment by itself is in tension with Planck at the 2.8σ level: this is significant since additive lensing systematics are more likely to impact the second moment than the third. So the third moment provides a useful check on the “low S_8” cosmic tension between the late time and early universe.

We expect to improve the analysis presented in this work and apply it to future data, such as at the final DES Y6 data. Based on the investigation performed in G20 we expect to further improve our constraining power on S_8 by roughly 20%, if we take into account the expected increase in the source number density. We plan to be able to model baryonic effects, which should allow us to push our analysis to smaller scales, improving constraints (up to 20%, Appendix A) and learning about baryonic physics. We are also planning to expand our modeling to include massive neutrinos and the full wCDM parameter space.

The simulated data used in this work has been generated using the public code FLASK (http://www.astro.iag.usp.br/~flask/), the public T17 simulations (http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/), and the public code PKDGRAV [72]. The full METACALIBRATION catalog will be made publicly available following publication, at the URL https://des.ncsa.illinois.edu/releases. The code used in this article will be shared on request to the corresponding author.

ACKNOWLEDGMENTS

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science
from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNpq Grant No. 465376/2014-2). This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

**APPENDIX A: SCALE CUTS**

We repeat on DES Y3 data the scale cut test we performed on simulated data in G20 in order to determine which part of the data vector can be used in the cosmological analysis. The reason the test is repeated is that some details of the analysis have been updated since G20 (mostly the nuisance parameters priors and the redshift distributions). The scale cut test is performed by contaminating a theory data vector with the known dominant systematic effect that is not part of our model; namely, baryonic feedback based on hydrodynamical simulations as described in G20. Then, we check that the cosmological parameters posterior obtained by analysing the contaminated data vector is not substantially biased with respect to the posterior with an uncontaminated data vector.

We adopted the “optimized scale cut criteria” for the ΛCDM cosmology adopted by the main DES cosmological analysis [15,18,108]. The criterion requires the peak of the marginalized 2D posterior of \( \Omega_\text{m} \) and \( S_8 \equiv \sigma_8(\Omega_\text{m}/0.3)^{0.5} \) obtained by analysing the contaminated data vector to be within 0.3 of the values obtained with the uncontaminated one. As we partially constrain \( n_s \), we also require the peak of the marginalized 2D posterior of \( n_s \) and \( S_8 \) to be within 0.5 of the baseline value. We arbitrarily chose a larger value for the \( n_s \) and \( S_8 \) criteria because \( n_s \) is only partially constrained and the posterior might be artificially too sharp. We also note that the DES Y3 3 × 2pt analysis does not assume any scale cut criteria on \( n_s \). We also check that the \( \chi^2 \) of the best-fitting cosmology of the analysis of the contaminated data vector is within 0.3 of the expected spread of the \( \chi^2 \) distribution. Since the length of the compressed data vector is 15, we require the best-fitting \( \chi^2 < 1.6^3 \). This second criterion ensures that the best-fitting \( \chi^2 \) from the analysis on data is unbiased. We note that these last two checks have not been included in the scale cut criteria in the main DES cosmological analysis.

In G20, we determined that a scale cut of \( R_0 = 24 h^{-1} \text{ Mpc} \) was sufficient (such that scales smaller than \( \theta_0 = R_0/\bar{\chi}(\bar{z}) \) were removed, where \( \bar{z} \) is the average of the mean redshift of different tomographic bins). When repeating this test, we had to use slightly large scales...
(\(R_0 = 28 h^{-1} \text{Mpc}\)) to pass the scale cut criteria, due to our updated analysis choices (e.g., inclusion of the shear-ratio likelihood, final values for redshift distributions, shape noise, effective number densities, covariance, etc.). Results are shown in Fig. 9; the peak of the 2D posterior of the contaminated data vector is 0.28σ off the baseline value in the \(\Omega_m-S_8\) plane, and 0.48σ in the \(n_s-S_8\) plane; we also obtain a best-fitting \(\chi^2 = 0.91 < 1.6\) for the contaminated data vector. Therefore, the scale cut of \(R_0 = 28 h^{-1} \text{Mpc}\) is deemed sufficient.

We note that our scales cut removes a significant number of data points from our measurement. This has a non-negligible impact on our constraining power. Using a simulated data vector, we estimate that we would improve our constraints on \(\Omega_0\) by a further 20% if we could apply no scale cut. This assumes we had a perfect knowledge of the baryonic effects on our data vector, which, unfortunately, is not the case for this analysis.

APPENDIX B: VALIDATION OF THE MODELING ON N-BODY SIMULATIONS

We repeat in this appendix the validation of our theoretical modeling performed in G20. We repeat that validation for two reasons: (1) some of our analysis choices have been updated (e.g., priors, redshift distributions, covariance, etc.); (2) we perform the validation on two independent N-body simulations (whereas in G20 we performed the validation on two sets of simulations). Results are shown in Fig. 9; the peak of the 2D marginalized constraints of the contaminated data vector; the filled square and circle show the peak of the contaminants for the contaminated and baseline data vectors, respectively.

We show first in Fig. 10 the comparison between theory predictions and the data vector as measured in the two sets of N-body simulations. For the data vector, we take the average of the data vector measured in every realization available. The mean offset between measurements and predictions is 0.5% and 8% for second and third moments of the T17 simulations, and 0.005% for both second and third moments of the PKDGRAV simulations (note that there are scale dependent residuals that are larger, but they average down when computing a mean offset). Note also that these numbers for the second moments are in agreement with the quoted uncertainties for the power spectrum for the two sets of simulations [71,72]; in particular, the pattern seen for the second moments of the PKDGRAV simulations is similar to that shown in [79].

We used the measured data vector (averaged over all the available realizations) and the scale cut determined in Appendix A to run two simulated cosmological analyses, one for each set of simulations. We compared the posteriors of \(\Omega_m\) and \(S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}\) with the posteriors obtained running the same cosmological analysis on a synthetic data vector at the “true” cosmology of the two simulations. Results are shown in Fig. 11, showing a good recovery of the true cosmological parameters.

APPENDIX C: ADDITIVE BIASES DUE TO PSF ERROR

We quantify in this appendix the level of contamination of our data vector due to additive biases related to PSF misestimation. PSF misestimation can cause additive biases in the measured galaxy shapes such that \(\gamma^\text{est} = \gamma + \delta e^\text{PSF} + \delta e^\text{noise}\). These spurious contributions can be characterized assuming a model for the PSF modeling errors and using a catalog of “reserved” stars that have not been used to train the PSF model. In what follows, we parametrize additive biases due to PSF misestimation following [14,114] (other modeling choices also exist in literature, e.g., [115]). In particular, we assume that

\[
\delta e^\text{PSF} = \alpha e_{\text{model}} + \beta (e_s - e_{\text{model}}) + \eta \left( e_s - T_s \right),
\]

\(083509-21\)
where $\alpha$, $\beta$, and $\eta$ are coefficients estimated from data, $e_s$ is the PSF ellipticity measured directly using the reserved stars catalog, $T_{s,\text{model}}$ is the modeled PSF size, and $T_s$ is the PSF size measured from the reserved stars catalog. The coefficients $\alpha$, $\beta$, and $\eta$ for the DES Y3 shape catalog have already been estimated in [14] for the nontomographic case and in [15] for the tomographic case. In what follows, we will use the values from [15], as we are interested in the contamination of our tomographic moments.

An empirical method was used to estimate the contribution to the measured moments due to PSF additive biases. We first created maps of $e_{s,\text{model}}$, $e_s$, and $T_s-T_{s,\text{model}}$ from the reserved stars catalog. Using the estimated values for $\alpha$, $\beta$, and $\eta$, we then created maps of $\delta e_{s,\text{model}}$, one for each tomographic bin. Last, we computed the second and third moments of the smoothed version of the $\delta e_{s,\text{model}}$ maps, in exactly the same way that we estimated the moments of the convergence maps (Sec. III B). In order to estimate the

![Graphical representation of the empirical method used to estimate the contribution to the measured moments due to PSF additive biases.](image-url)
contribution due to noise (that has to be subtracted from the raw, measured moments), we adopted a different technique as the two components of the $\delta e_{\text{sys}}$ field cannot just be randomly rotated as in the case of galaxies. We created two additional versions of the $\delta e_{\text{sys}}$ model maps, obtained by sampling two disjoint halves of the reserved stars catalog. We made sure the two halves spanned the footprint uniformly. We then measured the moments of the difference of the two maps, $\delta e_{\text{sys}}$ model; DIFF. In this way, the true signal should cancel, leaving only a contribution due to noise. The noise contribution to the moments of the $\delta e_{\text{sys}}$ model maps can be related to the signal measured from $\delta e_{\text{sys}}$ model; DIFF as follows:

$$h(\delta e_{\text{sys}} \text{ model; NOISE}) = 4 h(\delta e_{\text{sys}} \text{ model; DIFF})^2 i_{i,j},$$

$$h(\delta e_{\text{sys}} \text{ model; NOISE})^3 i_{i,j,k} = 8 h(\delta e_{\text{sys}} \text{ model; DIFF})^3 i_{i,j,k},$$

for any combination of tomographic bins $i$, $j$, and $k$. The second and third moments contribution due to PSF biases, once the noise term has been subtracted, is shown in Fig. 12. It can be seen clearly that such contribution is subdominant with respect to the moments of the convergence field, and that it mostly affects the large scales of the second moments. To further evaluate the impact of PSF modeling errors, we ran a cosmological analysis on a theory data vector contaminated by the measured moments of the PSF bias, and compared to the results obtained with a cosmological analysis performed on an uncontaminated theory data vector. The results are shown in Fig. 13, demonstrating that PSF additive biases have a negligible impact on our analysis.

APPENDIX D: NOISE TERMS AND SOURCE CLUSTERING

We can only have a noisy estimate of the shear field and so when computing the moments of the convergence maps the contribution due to noise has to be properly subtracted, as explained in Sec. III B [Eqs. (19) and (20)]. It is standard procedure to subtract only the contributions that are known to differ from zero, so as to not unnecessarily inflate the statistical uncertainty of our measurement. We show in Fig. 14 the noise terms as measured from the data. The noise contribution to the convergence map $\kappa_N$ has been obtained by randomly rotating the galaxy shapes and repeating the map-making procedure. For the second moments, we do not show the terms $h(\kappa_N)^{i,j}$ when $i = j$, as they are much larger than the measurement $h(\kappa_E)^{i,j}$ (at small scales, they are one order of magnitude larger) and are always subtracted. All the other terms are compatible with zero; the only exception concerns mixed terms of the form $h(\kappa\kappa_N)^{i,j,k}$, which presents some deviations from zero at small scales, especially in the moments involving the first tomographic bin (with a significance of $\chi^2 \sim 20–25/10$ d.o.f., depending on the bin combinations). This is in line with what we found in simulations in G20, where such terms did not vanish due to correlations between the pixels’ shape noise and the shear field value, induced by the intrinsic clustering of the sources [116]. These terms are subtracted from the measured moments before proceeding with the cosmological analysis; due to our scale cut this has a very small impact on the data vector used for the cosmological analysis. By using simulated data vectors with and without source clustering effects, we tested that this procedure is sufficient to remove the effect of source...
clustering and to have unbiased cosmological constraints (ignoring source clustering effects produces a shift in 2D $\Omega_m - S_8$ plane of only $0.08\sigma$).

**APPENDIX E: B MODES**

We show in this appendix the measured moments of the $B$-modes of the convergence maps. As we used the Kaiser-Squires algorithm to obtain the mass maps, non-null $B$-modes are expected as a consequence of mask effects [4], and are not necessarily associated with any observational systematic. The measured second moments are shown in Fig. 15. $B$-modes second moments are significantly nonzero; in the same figure, we also overplot the predicted $B$-modes given the best-fitting cosmology of the $E$-modes second moments, showing good agreement with the observed $B$-modes moments ($\chi^2 = 51/50$ d.o.f.). We do not detect any $B$-modes third moments at a significant level ($\chi^2 = 127/108$ d.o.f.); this is in line with the expected sensitivity of our data set and with the tests performed in G20.

**APPENDIX F: INTERNAL CONSISTENCY TESTS**

We quantify here the internal consistency of our data sets. Such tests, which rely on the PPD method, were performed prior to unblinding (using blinded data vectors) and were repeated after unblinding (although only the compatibility of the second and third moments was considered as an unblinding criterion).

**Compatibility between second and third moments.** This first test was one of the unblinding criteria. Using PPD, we can check that second and third moments posteriors are consistent with each other, so that we can run the analysis using the combined data vector. The PPD $p$-values for $p(\langle \kappa^2 \rangle | \langle \kappa^3 \rangle)$ and $p(\langle \kappa^3 \rangle | \langle \kappa^2 \rangle)$ are reported in Table IV and are well above the $p = 0.01$ threshold. We note that these values need not be the same as the two PPDs are not symmetric.

**Redshift tests.** Two types of internal consistency checks involving redshift distributions are performed. We performed these checks only using the combination of second and third moments; we did not perform them for second or third moments only.

The first check concerns the impact of removing individual redshift bins from the analysis. In order to perform this test, we again use the PPD. We first repeated our cosmological analysis removing all the second and third moments pairs and triples involving one particular redshift bin. We then sampled from those posteriors (one per bin), and compared using PPD to the observed second and third moments pairs and triples involving that particular redshift bin. This test is meant to highlight potential biases that...
of this alternative approach are shown in Fig. 16. We measure no significative difference with respect to our fiducial setup, demonstrating that for our analysis marginalizing over the uncertainties of the mean of the redshift distribution is sufficient.

Small scales vs. Large scales. We check for internal consistency between the small and large scales of our data vector. In order to do so, we split our data vector in two halves that have similar constraining power and that retain only either small or large scales. This is achieved by imposing a cut at the comoving scale of $56h^{-1}$ Mpc, which we converted to an angular scale cut as explained in Sec. IV. We use PPD to check for consistency between the two halves of the data vector. The PPD values are reported in Table IV; all the values are safely within our threshold. We note that for this test we considered the combination of second and third moments; we did not perform this test for second or third moments only. We show in Fig. 17 the posterior obtained using the two halves of the data vector; interestingly, the two halves of the data vector are associated to posteriors with a slightly different degeneracy direction in the $S_8-\Omega_m$ plane. This is a similar behaviour to what has been found in the DES Y3 cosmic shear analysis [15].

APPENDIX G: TESTS WITH ALTERNATIVE COVARIANCES

We explore in this appendix the effect on our posteriors of using different covariances. We have at our disposal three covariances, obtained using FLASK realizations, T17 N-body simulations, and PKDGRAV N-body simulations. These covariances assume different cosmologies, and in particular, model the higher-order moments of the convergence field slightly differently. This is easy to understand for FLASK covariance, since FLASK assumes the convergence field to be lognormal, which is only an approximation. We should also expect some differences between the T17 and PKDGRAV simulations, based on the different agreement with theory predictions shown in Fig. 11. Such differences probably stem from the different resolution settings of the two sets of N-body simulations.

When using the PKDGRAV and T17 covariances for the inference, we still use the FLASK covariance to compress the data vector since FLASK comes with the largest number of independent realizations. According to [117], this should not bias the inference, but (in the worst case) makes the compression suboptimal. We show in Fig. 18 the posteriors obtained using different covariance matrices.
FIG. 14. Measured moments involving the noise contribution to the convergence map. We do not show $(\kappa_N)^{i,j}$ for $i = j$, as these moments are much larger than the signal $(\kappa_E)^{i,j}$ and are always subtracted. The gray line is shown for reference and represents the expected theoretical signal for E-mode second and third moments. Gray shaded regions highlight the scales removed by the analysis.
for the combined second and third moments on data, and report the values of $S_8$ in Table II; the posteriors and the values of the constraints are very similar in the three cases, implying that our modeling of the covariance matrix is adequate.

APPENDIX H: PARAMETER 1D POSTERIORS AND TENSION WITH THE PRIORS

We test here whether the best-fitting models are in tension with their priors. This test was performed prior to unblinding. The 1D posteriors and their priors for all the parameters varied in this analysis are shown in Fig. 19. In order to quantify the level of tension with the priors we use the internal consistency test $p$-values. All internal consistency tests pass the predefined (arbitrary) threshold of 0.01. Besides the second vs. third moments tests, all the other tests have been performed on the data vector including the combination of second and third moments.

TABLE IV. Summary of internal consistency test $p$-values. All internal consistency tests pass the predefined (arbitrary) threshold of 0.01. Besides the second vs. third moments tests, all the other tests have been performed on the data vector including the combination of second and third moments.

| PPD test          | $p$-values |
|-------------------|------------|
| **Data splits**   |            |
| Bin 1 vs. no bin 1| 0.648      |
| Bin 2 vs. no bin 2| 0.148      |
| Bin 3 vs. no bin 3| 0.659      |
| Bin 4 vs. no bin 4| 0.260      |
| Large vs. small scales | 0.391 |
| Small vs. large scales | 0.350 |
| Second vs. third  | 0.32       |
| Third vs. second  | 0.49       |
a Gaussian estimator called the “update difference-in-mean” (UDM) statistic \[107\]. The UDM statistic compares the mean parameters from the prior \( \hat{\theta}_p \) with the updated values \( \hat{\theta}_p + d \) obtained running the analysis on data. In particular, we can define

$$ Q_{\text{UDM}} = (\hat{\theta}_p + d - \hat{\theta}_p)^T (C_p - C_p + d)^{-1} (\hat{\theta}_p + d - \hat{\theta}_p), $$

(H1)

where the difference in the mean of the parameters \( (\hat{\theta}_p + d - \hat{\theta}_p) \) is weighted by the inverse covariance of the parameters. If the parameters are Gaussian distributed then \( Q_{\text{UDM}} \) is chi-squared distributed with \( \text{rank}(C_p - C_p + d) \) degrees of freedom. The UDM tension for second moments, third moments, and the combination of second and third moments is 0.6\( \sigma \), 1.2\( \sigma \), and 0.8\( \sigma \), respectively, indicating no tension. We note that most of the parameter posteriors are actually prior dominated (without being in tension with the prior). This is fine as long as we trust our priors. The main parameters constrained by the analysis (\( \sigma_8 \) and \( \Omega_m \)) through \( S_8 \), and the intrinsic alignment amplitude \( A_{\text{IA}} \)), however, are not prior dominated. The effective number of parameters \( N_{\text{eff}} \) constrained by the analysis can be computed as

$$ N_{\text{eff}} = N - \text{tr}\left[(C_p)^{-1}(C_p + d)\right], $$

(H2)

where \( N \) is the number of free parameters in the analysis. For instance, if we restrict to the five cosmological parameters, \( N_{\text{eff}} \) is only 2.6, 1.5, and 2.6 for second moments, third moments, and the combination of second and third moments, respectively.

**APPENDIX I: COMPARISON BETWEEN SECOND MOMENTS AND COSMIC SHEAR WINDOW FUNCTION**

We discuss here a potential explanation for the \( \sim 1\sigma \) difference between the peaks of the DES Y3 cosmic shear and the second moments analysis in the \( S_8 - \Omega_m \) plane (Fig. 7). Both cosmic shear and second moments are Gaussian statistics and they both probe the shear power spectrum, but their posteriors do not have to perfectly overlap, as they weight power spectrum multipoles differently. Moreover, the process adopted to determine the scales that can be used for each tomographic bin is different for the two analyses; for the moments analysis we adopted a cut based on a well-determined physical scale \( \theta_0 = R_0/\chi(z) \) (Appendix A), whereas the cosmic shear analysis adopted a criterion that evenly distributed a given \( \Delta \chi^2 \) among tomographic bins (where the \( \Delta \chi^2 \) is computed between a synthetic data vector contaminated with baryonic effects and an uncontaminated data vector). Although both criteria are designed to minimize the impact of baryons on the \( S_8 - \Omega_m \) constraints, they can contribute to the different
FIG. 19. 1D marginalized posteriors for each parameter varied in the cosmological analysis. We show posteriors for second moments, third moments, and the combination of second and third moments, and compare them with their prior. Note that in some cases the edges of the 1D posteriors might look to extend over the prior edges; however, it is only a visual effect due to the smoothing of the plotting script close to the edge of the prior.
moments, $\xi$, multiplied by the window function $M$. The angular or smoothing scale allowed by the scale cut. The window statistics, the window function is computed at the minimum power spectrum by the mode-coupling matrix. For second moments, we also convolved the moments window function is given by Eq. (5). For second moments, we also convolved the power spectrum by the mode-coupling matrix.

![Shear power spectrum](image)

**Fig. 20.** Shear power spectrum for a given redshift bin multiplied by the window function $W$ of a given statistics (second moments, $\xi_{+-}$), normalized to their maximum value. For each statistics, the window function is computed at the minimum angular or smoothing scale allowed by the scale cut. The window functions for $\xi_{+-}$ are defined as $J_{0/4}(\ell\theta)$, whereas the second moments window function is $(2\ell + 1)[W_\ell(\theta_0)]^2$, with $W_\ell(\theta_0)$ given by Eq. (5). For second moments, we also convolved the power spectrum by the mode-coupling matrix.

sensitivity of the statistics to the shear power spectrum multipoles.

Figure 20 shows, for the first and the last redshift bins, how the different statistics weight the multipoles of the shear power spectrum at the minimum angular scale allowed by their scale cut. The cosmic shear scale cut allows $\xi_{+-}$ to probe significantly higher multipoles compared to the second moments, whose window function is more compact. When considering the redshift bin 1 (4), $\sim 30$ ($\sim 25$) per cent of the S/N of the $\xi_{+-}$ data vector passing the scale cut comes from $\ell' > 200$. For second moments the contribution to the S/N from $\ell' > 200$ is significantly smaller; $\sim 1$ ($\sim 10$) per cent. Although a more quantitative assessment of the compatibility between the second moments and cosmic shear constraints should be performed via PPD, we consider the pieces of evidence provided in this appendix sufficient to justify the differences between the two analyses shown in Fig. 7.

[1] A. Einstein, *Science* **84**, 506 (1936).
[2] C. Chang, A. Pujol, B. Mawdsley et al., *Mon. Not. R. Astron. Soc.* **475**, 3165 (2018).
[3] C. Chang, V. Vikram, B. Jain et al., *Phys. Rev. Lett.* **115**, 051301 (2015).
[4] N. Jeffrey, M. Gatti, C. Chang et al., *Mon. Not. R. Astron. Soc.* **505**, 4626 (2021).
[5] J. Liu, A. Petri, Z. Haiman, L. Hui, J. M. Kratochvil, and M. May, *Phys. Rev. D* **91**, 063507 (2015).
[6] M. Oguri, S. Miyazaki, C. Hikage et al., *Publ. Astron. Soc. Jpn.* **70**, S26 (2018).
[7] L. Van Waerbeke, J. Benjamin, T. Erben et al., *Mon. Not. R. Astron. Soc.* **433**, 3373 (2013).
[8] V. Vikram, C. Chang, B. Jain et al., *Phys. Rev. D* **92**, 022006 (2015).
[9] Dark Energy Survey Collaboration, *Mon. Not. R. Astron. Soc.* **460**, 1270 (2016).
[10] K. Kuijken, C. Heymans, H. Hildebrandt et al., *Mon. Not. R. Astron. Soc.* **454**, 3500 (2015).
[11] H. Aihara, N. Arimoto, R. Armstrong et al., *Publ. Astron. Soc. Jpn.* **70**, S4 (2018).
[12] P. A. Abell, J. Allison et al. (LSST Science Collaboration), arXiv:0912.0201.
[13] R. Laureijs, J. Amiaux, S. Arduri et al., arXiv:1110.3193.
[14] M. Gatti, E. Sheldon, A. Amon et al., *Mon. Not. R. Astron. Soc.* **504**, 4312 (2021).
[15] A. Amon, D. Gruen, M. A. Troxel et al., *Phys. Rev. D* **105**, 023514 (2022).
[16] C. Hikage, M. Oguri, T. Hamana et al., *Publ. Astron. Soc. Jpn.* **71**, 43 (2019).
[17] H. Hildebrandt, M. Viola, C. Heymans et al., *Mon. Not. R. Astron. Soc.* **465**, 1454 (2017).
[18] L. F. Secco, S. Samuroff, E. Krause et al., *Phys. Rev. D* **105**, 023515 (2022).
[19] M. A. Troxel, N. MacCrann, and J. Zuntz et al., *Phys. Rev. D* **98**, 043528 (2018).
[20] T. Hamana, M. Shirasaki, S. Miyazaki et al., *Publ. Astron. Soc. Jpn.* **72**, 16 (2020).
[21] M. Asgari, C.-A. Lin, B. Joachimi et al., *Astron. Astrophys.* **645**, A104 (2021).
[22] P. Coles and B. Jones, *Mon. Not. R. Astron. Soc.* **248**, 1 (1991).
[23] E. Hubble, *Astrophys. J.* **79**, 8 (1934).
[24] V. Wild, J. A. Peacock, O. Lahav et al., *Mon. Not. R. Astron. Soc.* **356**, 247 (2005).
[25] L. Clerkin, D. Kirk, M. Manera et al., *Mon. Not. R. Astron. Soc.* **466**, 1444 (2017).
[26] M. Gatti, C. Chang, O. Friedrich et al., *Mon. Not. R. Astron. Soc.* **498**, 4060 (2020).
[27] A. Petri, J. Liu, Z. Haiman, M. May, L. Hui, and J. M. Kratochvil, *Phys. Rev. D* **91**, 103511 (2015).
[28] S. Vafaei, T. Lu, L. van Waerbeke, E. Semboloni, C. Heymans, and U.-L. Pen, *Astropart. Phys.* **32**, 340 (2010).
[29] D. Zürcher, J. Fluri, R. Sgier, T. Kacprzak, and A. Refregier, *J. Cosmol. Astropart. Phys.* 01(2021)028.
[30] V. F. Cardone, S. Camera, R. Mainini, A. Romano, A. Diaferio, R. Madau, and R. Scaramella, *Mon. Not. R. Astron. Soc.* **430**, 2896 (2013).
[31] A. Peel, V. Pettorino, C. Giocoli, J.-L. Starck, and M. Baldi, *Astron. Astrophys.* **619**, A38 (2018).
[32] V. Ajani, A. Peel, V. Pettorino, J.-L. Starck, Z. Li, and J. Liu, *Phys. Rev. D* **102**, 103531 (2020).
[33] J. P. Dietrich and J. Hartlap, *Mon. Not. R. Astron. Soc.* **402**, 1049 (2010).
DARK ENERGY SURVEY YEAR 3 RESULTS: COSMOLOGY WITH 

[34] T. Kacprzak, D. Kirk, O. Friedrich et al., Mon. Not. R. Astron. Soc. 463, 3653 (2016).

[35] J. M. Kratochvil, Z. Haiman, and M. May, Phys. Rev. D 81, 043519 (2010).

[36] N. Martinet, P. Schneider, H. Hildebrandt et al., Mon. Not. R. Astron. Soc. 474, 712 (2018).

[37] H. Shan, X. Liu, H. Hildebrandt et al., Mon. Not. R. Astron. Soc. 474, 1116 (2018).

[38] M. Vicinanza, V. F. Cardone, R. Maoli, R. Scaramella, and X. Er, arXiv:1606.03892.

[39] M. Vicinanza, V. F. Cardone, R. Maoli, R. Scaramella, and X. Er, Phys. Rev. D 97, 023519 (2018).

[40] L. Fu, M. Kilbinger, T. Erben et al., Mon. Not. R. Astron. Soc. 441, 2725 (2014).

[41] E. Sembolini, T. Schrabback, L. van Waerbeke, S. Vafaei, J. Hartlap, and S. Hilbert, Mon. Not. R. Astron. Soc. 410, 143 (2011).

[42] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. 344, 857 (2003).

[43] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. 348, 897 (2004).

[44] J. M. Kratochvil, E. A. Lim, S. Wang, Z. Haiman, M. May, and K. Huffenberger, Phys. Rev. D 85, 103513 (2012).

[45] C. Parroni, V. F. Cardone, R. Maoli, and R. Scaramella, Mon. Not. R. Astron. Soc. 633, A71 (2020).

[46] M. Vicinanza, V. F. Cardone, R. Maoli, R. Scaramella, X. Er, and T. Tereno, Phys. Rev. D 99, 043534 (2019).

[47] J. Fluri, T. Kacprzak, A. Lucchi, A. Refregier, A. Amara, T. Hofmann, and A. Schneider, Phys. Rev. D 100, 063514 (2019).

[48] J. Fluri, T. Kacprzak, A. Refregier, A. Amara, A. Lucchi, and T. Hofmann, Phys. Rev. D 98, 123518 (2018).

[49] N. Jeffrey, J. Alsing, and F. Lanusse, Mon. Not. R. Astron. Soc. 501, 954 (2020).

[50] D. Bibi, B. Á. Pataki, and I. Csabai, Nat. Astron. 3, 93 (2019).

[51] P. Fosalba, E. Gaztanaga, F. J. Castander, and M. Manera, Mon. Not. R. Astron. Soc. 391, 435 (2008).

[52] E. Gaztanaga and F. Bernardeau, Astron. Astrophys. 331, 829 (1998).

[53] B. Jain and U. Seljak, Astrophys. J. 484, 560 (1997).

[54] A. Pujol, C. Chang, E. Gaztanaga et al., Mon. Not. R. Astron. Soc. 462, 35 (2016).

[55] B. Flaugher, H. T. Diehl, K. Honscheid et al., Astron. J. 150, 150 (2015).

[56] Dark Energy Survey Collaboration, Astrophys. J. Suppl. Ser. 239, 18 (2018).

[57] E. Morganson, R. A. Gruendl, F. Menanteau et al., Publ. Astron. Soc. Pac. 130, 074501 (2018).

[58] I. Sevilla et al., arXiv:1106.6741.

[59] I. Sevilla-Noarbe, K. Bechtol, M. Carrasco Kind et al., Astrophys. J. Suppl. Ser. 254, 24 (2021).

[60] E. Huff and R. Mandelbaum, arXiv:1702.02600.

[61] E. S. Sheldon and E. M. Huff, Astrophys. J. 841, 24 (2017).

[62] J. Zuntz, E. Sheldon et al., Mon. Not. R. Astron. Soc. 481, 1149 (2018).

[63] N. MacCrann, M. R. Becker, J. McCullough et al., Mon. Not. R. Astron. Soc. 509, 3371 (2021).

[64] E. S. Sheldon, M. R. Becker, N. MacCrann, and M. Jarvis, Astrophys. J. 902, 138 (2020).

[65] J. Myles, A. Alarcon, and A. Amon et al., Mon. Not. R. Astron. Soc. 505, 4249 (2021).

[66] M. Gatti, G. Giannini, G. M. Bernstein et al., Mon. Not. R. Astron. Soc. 510, 1223 (2021).

[67] C. Sánchez, J. Prat, G. Zacharek et al., Phys. Rev. D 105, 083529 (2022).

[68] J. Muir, G. M. Bernstein, D. Huterer et al., Mon. Not. R. Astron. Soc. 494, 4454 (2020).

[69] T. M. C. Abbott, F. B. Abdalla, A. Alarcon et al., Phys. Rev. D 98, 043526 (2018).

[70] H. S. Xavier, F. B. Abdalla, and B. Joachimi, Mon. Not. R. Astron. Soc. 459, 3693 (2016).

[71] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, Astrophys. J. 850, 24 (2017).

[72] D. Poter, J. Stadel, and R. Teyssier, Comput. Astrophys. Cosmol. 4, 2 (2017).

[73] O. Friedrich, F. Andrade-Oliveira, H. Camacho et al., Mon. Not. R. Astron. Soc. 508, 3125 (2021).

[74] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. 622, 759 (2005).

[75] V. Springel, Mon. Not. R. Astron. Soc. 364, 1105 (2005).

[76] M. Crocce, S. Pueblas, and R. Scoccimarro, Mon. Not. R. Astron. Soc. 373, 369 (2006).

[77] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, Astrophys. J. 761, 152 (2012).

[78] T. Hamana, J. Sakurai, M. Koike, and L. Miller, Publ. Astron. Soc. Jpn. 67, 34 (2015).

[79] D. Zürcher et al., Mon. Not. R. Astron. Soc. 511, 2075 (2022).

[80] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. 367, 1 (2002).

[81] R. Scoccimarro and H. M. P. Couchman, Mon. Not. R. Astron. Soc. 325, 1312 (2001).

[82] L. Van Waerbeke, T. Hamana, R. Scoccimarro, S. Colombi, and F. Bernardeau, Mon. Not. R. Astron. Soc. 322, 918 (2001).

[83] M. L. Brown, P. G. Castro, and A. N. Taylor, Mon. Not. R. Astron. Soc. 360, 1262 (2005).

[84] C. Hikage, M. Takada, T. Hamana, and D. Spergel, Mon. Not. R. Astron. Soc. 412, 65 (2011).

[85] R. Takahashi, S. Sonn, M. Takada, and I. Kayo, Mon. Not. R. Astron. Soc. 444, 3473 (2014).

[86] J. P. Cordero, J. Harrison, R. P. Rollins et al., Mon. Not. R. Astron. Soc. 511, 2170 (2022).

[87] S. Bridle and L. King, New J. Phys. 9, 444 (2007).

[88] C. M. Hirata and U. Seljak, Phys. Rev. D 70, 063526 (2004).

[89] S. Pyne, A. Tenneti, and B. Joachimi, arXiv:2204.10342.

[90] A. J. S. Hamilton, Mon. Not. R. Astron. Soc. 322, 419 (2001).

[91] J. A. Blazek, N. MacCrann, M. A. Troxel, and X. Fang, Phys. Rev. D 100, 103506 (2019).

[92] A. Porredon, M. Croce, P. Fosalba et al., Phys. Rev. D 103, 043503 (2021).
[93] A. Porredon et al., arXiv:2105.13546 [Phys. Rev. D (to be published)].
[94] N. Kaiser and G. Squires, Astrophys. J. 404, 441 (1993).
[95] C. G. R. Wallis, J. D. McEwen, T. D. Kitching, B. Leistedt, and A. Plouviez, massmappy: Mapping dark matter on the celestial sphere, https://ui.adsabs.harvard.edu/abs/2022ascl.soft07035W.
[96] D. Gualdi, M. Manera, B. Joachimi, and O. Lahav, Mon. Not. R. Astron. Soc. 476, 4045 (2018).
[97] A. F. Heavens, R. Jimenez, and O. Lahav, Mon. Not. R. Astron. Soc. 317, 965 (2000).
[98] M. Tegmark, A. N. Taylor, and A. F. Heavens, Astrophys. J. 480, 22 (1997).
[99] J. Hartlap, P. Simon, and P. Schneider, Astron. Astrophys. 464, 399 (2007).
[100] S. Dodelson and M. D. Schneider, Phys. Rev. D 88, 063537 (2013).
[101] O. Friedrich and T. Eifler, Mon. Not. R. Astron. Soc. 473, 4150 (2018).
[102] W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. R. Astron. Soc. 450, L61 (2015).
[103] W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. R. Astron. Soc. 453, 4385 (2015).
[104] S. Habib, K. Heitmann, D. Higdon, C. Nakhleh, and B. Williams, Phys. Rev. D 76, 083503 (2007).
[105] K. Heitmann, D. Higdon, C. Nakhleh, and S. Habib, Astrophys. J. 646, L1 (2006).
[106] C. Doux, E. Baxter, P. Lemos et al., Mon. Not. R. Astron. Soc. 503, 2688 (2021).
[107] M. Raveri and W. Hu, Phys. Rev. D 99, 043506 (2019).
[108] T. M. C. Abbott et al. (DES Collaboration), Phys. Rev. D 105, 023520 (2022).
[109] C. Heymans, T. Tröster, M. Asgari et al., Astron. Astrophys. 646, A140 (2021).
[110] T. Hamana, M. Shirasaki, S. Miyazaki et al., Publ. Astron. Soc. Jpn. 72, 16 (2020).
[111] N. Aghanim, Y. Akrami, M. Ashdown et al., Astron. Astrophys. 641, A6 (2020).
[112] M. Raveri, G. Zacharegkas, and W. Hu, Phys. Rev. D 101, 103527 (2020).
[113] C. Chang, M. Wang, S. Dodelson et al., Mon. Not. R. Astron. Soc. 482, 3696 (2019).
[114] M. Jarvis, E. Sheldon, J. Zuntz et al., Mon. Not. R. Astron. Soc. 460, 2245 (2016).
[115] B. Giblin, C. Heymans, M. Asgari et al., Astron. Astrophys. 645, A105 (2021).
[116] E. Krause et al., arXiv:2105.13548 [Mon. Not. R. Astron. Soc. (to be published)].
[117] A. F. Heavens, E. Sellentin, D. de Mijolla, and A. Vianello, Mon. Not. R. Astron. Soc. 472, 4244 (2017).