

# Dark Energy Survey Year 3 results: simulation-based cosmological inference with wavelet harmonics, scattering transforms, and moments of weak lensing mass maps II. Cosmological results

M. Gatti,<sup>1,\*</sup> G. Campailla,<sup>2</sup> N. Jeffrey,<sup>3</sup> L. Whiteway,<sup>3</sup> A. Porredon,<sup>4</sup> J. Prat,<sup>5</sup> J. Williamson,<sup>3</sup> M. Raveri,<sup>2</sup> B. Jain,<sup>1</sup> V. Ajani,<sup>6,7</sup> G. Giannini,<sup>5</sup> M. Yamamoto,<sup>8</sup> C. Zhou,<sup>9,7</sup> J. Blazek,<sup>10</sup> D. Anbajagane,<sup>5</sup> S. Samuroff,<sup>10</sup> T. Kacprzak,<sup>6</sup> A. Alarcon,<sup>11</sup> A. Amon,<sup>12,13</sup> K. Bechtol,<sup>14</sup> M. Becker,<sup>15</sup> G. Bernstein,<sup>1</sup> A. Campos,<sup>14</sup> C. Chang,<sup>5,7</sup> R. Chen,<sup>8</sup> A. Choi,<sup>16</sup> C. Davis,<sup>17</sup> J. Derose,<sup>18</sup> H. T. Diehl,<sup>19</sup> S. Dodelson,<sup>20,21</sup> C. Doux,<sup>22</sup> K. Eckert,<sup>1</sup> J. Elvin-Poole,<sup>23</sup> S. Everett,<sup>24</sup> A. Ferte,<sup>25</sup> D. Gruen,<sup>26</sup> R. Gruendl,<sup>27,28</sup> I. Harrison,<sup>29</sup> W. G. Hartley,<sup>30</sup> K. Herner,<sup>19</sup> E. M. Huff,<sup>24</sup> M. Jarvis,<sup>1</sup> N. Kuropatkin,<sup>19</sup> P. F. Leget,<sup>17</sup> N. MacCrann,<sup>31</sup> J. McCullough,<sup>17</sup> J. Myles,<sup>32</sup> A. Navarro-Alsina,<sup>33</sup> S. Pandey,<sup>1</sup> R. P. Rollins,<sup>34</sup> A. Roodman,<sup>17,25</sup> C. Sanchez,<sup>1</sup> L. F. Secco,<sup>5</sup> I. Sevilla-Noarbe,<sup>35</sup> E. Sheldon,<sup>36</sup> T. Shin,<sup>37</sup> M. Troxel,<sup>8</sup> I. Tutusaus,<sup>38</sup> T. N. Varga,<sup>39,40,41</sup> B. Yanny,<sup>19</sup> B. Yin,<sup>20</sup> Y. Zhang,<sup>42</sup> J. Zuntz,<sup>43</sup> T. M. C. Abbott,<sup>42</sup> M. Aguena,<sup>44</sup> S. S. Allam,<sup>19</sup> O. Alves,<sup>45</sup> F. Andrade-Oliveira,<sup>45</sup> D. Bacon,<sup>46</sup> S. Bocquet,<sup>26</sup> D. Brooks,<sup>3</sup> A. Carnero Rosell,<sup>47,44</sup> J. Carretero,<sup>48</sup> L. N. da Costa,<sup>44</sup> M. E. S. Pereira,<sup>49</sup> J. De Vicente,<sup>35</sup> I. Ferrero,<sup>50</sup> J. Frieman,<sup>19,5</sup> J. García-Bellido,<sup>51</sup> E. Gaztanaga,<sup>52,46,53</sup> G. Gutierrez,<sup>19</sup> S. R. Hinton,<sup>54</sup> D. L. Hollowood,<sup>55</sup> K. Honscheid,<sup>56,57</sup> D. J. James,<sup>58</sup> K. Kuehn,<sup>59,60</sup> O. Lahav,<sup>3</sup> S. Lee,<sup>24</sup> J. L. Marshall,<sup>61</sup> J. Mena-Fernández,<sup>62</sup> R. Miquel,<sup>63,48</sup> A. Pieres,<sup>44,64</sup> A. A. Plazas Malagón,<sup>17,25</sup> E. Sanchez,<sup>35</sup> D. Sanchez Cid,<sup>35</sup> M. Schubnell,<sup>45</sup> M. Smith,<sup>65</sup> E. Suchyta,<sup>66</sup> G. Tarle,<sup>45</sup> N. Weaverdyck,<sup>67,68</sup> J. Weller,<sup>40,41</sup> and P. Wiseman<sup>65</sup>

<sup>1</sup>Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>2</sup>Department of Physics, University of Genova and INFN, Via Dodecaneso 33, 16146, Genova, Italy

<sup>3</sup>Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

<sup>4</sup>Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing, 44780 Bochum, Germany

<sup>5</sup>Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

<sup>6</sup>Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland

<sup>7</sup>Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA

<sup>8</sup>Department of Physics, Duke University Durham, NC 27708, USA

<sup>9</sup>Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA

<sup>10</sup>Department of Physics, Northeastern University, Boston, MA 02115, USA

<sup>11</sup>Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

<sup>12</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>13</sup>Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>14</sup>Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, WI 53706-1390, USA

<sup>15</sup>Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

<sup>16</sup>NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

<sup>17</sup>Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

<sup>18</sup>Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

<sup>19</sup>Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

<sup>20</sup>Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15312, USA

<sup>21</sup>NSF AI Planning Institute for Physics of the Future, Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>22</sup>Université Grenoble Alpes, CNRS, LPSC-IN2P3, 38000 Grenoble, France

<sup>23</sup>Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

<sup>24</sup>Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

<sup>25</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

<sup>26</sup>University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany

<sup>27</sup>Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

<sup>28</sup>Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA

<sup>29</sup>School of Physics and Astronomy, Cardiff University, CF24 3AA, UK

<sup>30</sup>Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland

<sup>31</sup>Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK

<sup>32</sup>Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA

<sup>33</sup>Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil

- <sup>34</sup>*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*
- <sup>35</sup>*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
- <sup>36</sup>*Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA*
- <sup>37</sup>*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA*
- <sup>38</sup>*Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France*
- <sup>39</sup>*Excellence Cluster Origins, Boltzmannstr. 2, 85748 Garching, Germany*
- <sup>40</sup>*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany*
- <sup>41</sup>*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany*
- <sup>42</sup>*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
- <sup>43</sup>*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*
- <sup>44</sup>*Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
- <sup>45</sup>*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
- <sup>46</sup>*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*
- <sup>47</sup>*Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain*
- <sup>48</sup>*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain*
- <sup>49</sup>*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany*
- <sup>50</sup>*Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway*
- <sup>51</sup>*Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
- <sup>52</sup>*Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain*
- <sup>53</sup>*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*
- <sup>54</sup>*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
- <sup>55</sup>*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
- <sup>56</sup>*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
- <sup>57</sup>*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
- <sup>58</sup>*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
- <sup>59</sup>*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
- <sup>60</sup>*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
- <sup>61</sup>*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*
- <sup>62</sup>*LPSC Grenoble - 53, Avenue des Martyrs 38026 Grenoble, France*
- <sup>63</sup>*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
- <sup>64</sup>*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
- <sup>65</sup>*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*
- <sup>66</sup>*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831*
- <sup>67</sup>*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*
- <sup>68</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

We present a simulation-based cosmological analysis using a combination of Gaussian and non-Gaussian statistics of the weak lensing mass (convergence) maps from the first three years (Y3) of the Dark Energy Survey (DES). We implement: 1) second and third moments; 2) wavelet phase harmonics; 3) the scattering transform. Our analysis is fully based on simulations, spans a space of seven  $\Lambda$ CDM cosmological parameters, and forward models the most relevant sources of systematics inherent in the data: masks, noise variations, clustering of the sources, intrinsic alignments, and shear and redshift calibration. We implement a neural network compression of the summary statistics, and we estimate the parameter posteriors using a simulation-based inference approach. Including and combining different non-Gaussian statistics is a powerful tool that strongly improves constraints over Gaussian statistics (in our case, the second moments); in particular, the Figure of Merit  $\text{FoM}(S_8, \Omega_m)$  is improved by 70 percent ( $\Lambda$ CDM) and 90 percent ( $w$ CDM). When all the summary statistics are combined, we achieve a 2 percent constraint on the amplitude of fluctuations parameter  $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ , obtaining  $S_8 = 0.794 \pm 0.017$  ( $\Lambda$ CDM) and  $S_8 = 0.817 \pm 0.021$  ( $w$ CDM), and a  $\sim 10$  percent constraint on  $\Omega_m$ , obtaining  $\Omega_m = 0.259 \pm 0.025$  ( $\Lambda$ CDM) and  $\Omega_m = 0.273 \pm 0.029$  ( $w$ CDM). In the context of the  $w$ CDM scenario, these statistics also strengthen the constraints on the parameter  $w$ , obtaining  $w < -0.72$ . The constraints from different statistics are shown to be internally consistent (with a  $p$ -value  $> 0.1$  for all combinations of statistics examined). We compare our results to other weak lensing results from the DES Y3 data, finding good consistency; we also compare with results

from external datasets, such as PLANCK constraints from the Cosmic Microwave Background, finding statistical agreement, with discrepancies no greater than  $< 2.2\sigma$ .

## I. INTRODUCTION

Weak gravitational lensing serves as an efficient technique for investigating the large-scale structure and matter distribution throughout the Universe. This method enables us to infer the distribution of matter lying in the foreground by examining the minor distortions evident in the shapes of background galaxies. Such effects of weak gravitational lensing provide a window into the history of the Universe’s expansion and its geometric properties, acting as tools for understanding the underlying structure of the cosmos.

Conventional methods of cosmological analysis employ Gaussian statistics, using tools such as two-point correlation functions or power spectra to analyze the lensing signals. The most up-to-date analyses of this type are from the Dark Energy Survey (DES, Amon et al. 6, Secco et al. 76), the Kilo-Degree Survey (KiDS, Asgari et al. 9, Li et al. 50), and Hyper Suprime-Cam (HSC, Dalal et al. 22, Li et al. 51). A growing number of studies, however, have investigated and highlighted the role of non-Gaussian statistics in improving cosmological constraints, as the lensing observables carry information beyond that probed by standard Gaussian statistics. Examples of non-Gaussian statistics investigated include higher-order moments [17, 29, 31, 64, 65, 68, 83–85], peak counts [2, 23, 36, 46, 48, 52, 58, 64, 77, 87, 88], one-point probability distributions [11, 14, 79], Minkowski functionals [49, 62, 65, 86], Betti numbers [25, 63], persistent homology [38, 39], scattering transform coefficients [18, 19, 80, 81], wavelet phase harmonic moments [4], kNN and CDFs [7, 10], map-level inference [13, 67], and machine-learning methods [26, 27, 41, 53, 72].

This work uses a comprehensive simulation-based inference approach to analyze the weak lensing data from the first three years of the Dark Energy Survey (DES Y3). Our analysis concentrates on three distinct statistics: second and third-order moments, wavelet phase harmonic moments, and the scattering transform. While the second and third moments have been previously applied to DES Y3 data [31], their implementation was based on a theoretical framework for the modelling of the moments. In contrast, our current approach is driven by simulations; i.e., we use simulations to model our observables. Moreover, this research is the first application of wavelet phase harmonic moments to weak lensing data, and one of the first applications of scattering transform (see [18]). These two statistics have become prominent for their similarities to convolutional neural networks (CNNs), designed to extract information from fields in a manner comparable to CNNs [57], but without the requirement for training. In [32] (hereafter paper

I), we showed that these non-Gaussian statistics provide additional information beyond that of third moments, indicating their potential utility.

We use the *Gower Street simulations* (Jeffrey et al. 44), a suite of N-body simulations spanning a seven-dimensional  $w$ CDM parameter space. We incorporate key sources of systematic uncertainties in our forward model such as photometric redshift uncertainties, shear calibration errors, intrinsic alignments, and the effects of source clustering [33]. We implement an efficient neural network compression of our summary statistics, and we estimate the parameter posteriors via neural density estimation of the likelihood surface, using flexible neural networks that do not imposing restrictive assumptions about the likelihood or data model. This paper builds upon a companion work (paper I), in which the methodology was validated using simulations and a series of systematic observational tests were conducted to verify the robustness of the analysis. This work applies the validated methodology to the DES Y3 data to yield cosmological constraints; we also compare these findings with results from other DES Y3 analyses and from external datasets.

A companion DES analysis (Jeffrey et al. 44) uses the same simulation-based inference pipeline as described in this paper, but using convolutional neural networks (map-level inference), peak counts, and power spectra to infer cosmology. Their results are consistent with the results presented in this work. All of these are analyses of DES Y3 data, pending the final full DES Year 6 data.

This paper is organized as follows: Section 2 summarizes the survey data as well as the simulations used for our model predictions and validation. Section 3 briefly describes the summary statistics used in this paper and the data compression. Section 4 briefly describes our simulation-based inference pipeline for parameter inference. Section 5 presents the pre-unblinding tests, and Section 6 discusses our main results, together with comparisons against other DES results and external probes.

## II. DATA AND SIMULATIONS

### A. DES Y3 weak lensing catalogue and weak lensing mass maps

We use the DES Y3 weak lensing catalogue for our analysis [34]. This extensive catalogue comprises 100,204,026 galaxies, offering a weighted effective galaxy number density of 5.59 galaxies per square arcminute across a 4139 square degree area. The catalogue is based on the METACALIBRATION algorithm [40, 78], which calculates self-calibrated shear measurements from the multi-band, noisy images of observed objects. To correct for any residual calibration issues such as multiplicative shear bias, we incorporate adjustments based

---

\* marcogatti29@gmail.com

Table I. Model parameters (first column), their distribution in the Gower Street sims and in the mock catalogues derived from these sims (second column), and the prior used in the cosmological analysis (third column, shown only where different from the second column). The analysis prior can differ from the distribution of the samples as long as these parameters have been explicitly used during the training of the Neural Density Estimators (NDEs) when learning the likelihood surface; see §IV for more details. \* In our simulation runs we usually excluded values of  $w$  less than  $-1$ , but 64 simulations were run without this constraint. These were still used to train our NDEs, although we applied a strict prior of  $w > -1$  for the analysis.

Parameter	Mocks parameters distribution	Analysis prior
$\Omega_m$	mixed active-learning in $\mathcal{U}(0.15, 0.52)$	$\mathcal{U}(0.15, 0.52)$
$S_8$	mixed active-learning in $\mathcal{U}(0.5, 1.0)$	$\mathcal{U}(0.5, 1.0)$
$w$	$\mathcal{N}(-1, \frac{1}{3})$ for $-1 < w < -\frac{1}{3}$ 0 else*	$\mathcal{U}(-1, -\frac{1}{3})$
$n_s$	$\mathcal{N}(0.9649, 0.0063)$	
$h$	$\mathcal{N}(0.7022, 0.0245)$	
$\Omega_b h^2$	$\mathcal{N}(0.02237, 0.00015)$	
$m_\nu$	$\exp(\mathcal{U}[\log(0.06), \log(0.14)])$	
$A_{IA}$	$\mathcal{U}[-3, 3]$	
$\eta_{IA}$	$\mathcal{U}[-5, 5]$	
$m_1$	$\mathcal{N}(-0.0063, 0.0091)$	
$m_2$	$\mathcal{N}(-0.0198, 0.0078)$	
$m_3$	$\mathcal{N}(-0.0241, 0.0076)$	
$m_4$	$\mathcal{N}(-0.0369, 0.0076)$	
$\bar{n}_i(z)$	$P_{\text{HYPERRANK}}(\bar{n}_i(z) x_{\text{phot}})$	

on sophisticated image simulations [54]. The catalogue includes a per-galaxy inverse variance weight, which enhances the signal-to-noise ratio of our measurements. The galaxies are partitioned into four tomographic bins of roughly equal number density, as described in [60].

The redshift distributions for these bins are derived using the SOMPZ method [60], augmented by clustering-based redshift information [30] and a correction for the redshift-dependent influences of blending [54]. We produce weak lensing mass maps for each tomographic bin using a full-sky extension of the Kaiser and Squires algorithm [47], [42]. The maps are pixelized using HEALPIX [35] at a resolution of NSIDE = 512; this yields a pixel size of approximately 6.9 arcminutes.

## B. Simulations

Here is a brief summary of the simulations used in this work; see paper I for details.

The Gower Street simulation suite [44] is key to our inference process. The mocks created from the Gower Street suite serve two purposes: they are used both for compressing the summary statistics and for performing the cosmological inference. The suite consists of 791

gravity-only full-sky N-body simulations, produced using the PKDGRAV3 code [69]. The simulations span a seven-dimensional parameter space in  $w$ CDM ( $\Omega_m, \sigma_8, n_s, h \equiv h_{100}, \Omega_b, w, m_\nu$ ); see Table I. This table uses the parameter  $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ .

Each full-sky simulation can be divided into four non-overlapping DES sky footprints, yielding 3,164 independent mock DES surveys. To generate additional pseudo-independent DES Y3 shear mock maps, we rotated the four independent DES Y3 footprints by 45, 90, and 135 degrees in galactic longitude, thereby covering different (but overlapping) regions of the full-sky map. For each map we generated two distinct noise realizations, finally yielding a total of 25,312 pseudo-independent noisy DES Y3 mock maps. The maps from the first noise realization are used for the compression step, while those from the second noise realization are used for the cosmological inference step.

For this analysis, we augmented the Gower Street simulation suite with N-body simulations from the **DarkGridV1** suite [88, 89]. This integration was done because, unlike paper I, we include  $\Lambda$ CDM results in this work, and the **DarkGridV1** simulations are exclusively  $\Lambda$ CDM simulations, in contrast to those from Gower Street. We later discover, however, that this addition has a negligible impact on our constraints (see Appendix D), and the Gower Street simulations were sufficient also for the  $\Lambda$ CDM case. The **DarkGridV1** simulations suite explores 58 different  $\Lambda$ CDM cosmologies, varying  $\Omega_m$  and  $\sigma_8$ . Each cosmology is represented by five independent full-sky simulations. For each full-sky simulation, we applied the same procedure as used for the Gower Street suite to generate multiple DES Y3 mock catalogues, resulting in an additional 2,320 mocks. These mocks are added to those from the Gower Street suite and are used both in the compression of the summary statistics and in the cosmological inference.

In this analysis we only explicitly learn the likelihood surface for the cosmological parameters  $\Omega_m, S_8, w$ , and the intrinsic alignment amplitude  $A_{IA}$ ; the dependence on other parameters is not explicitly learned, but is effectively marginalised over according to the distribution followed by the simulations (§IV). In contrast to the Gower Street suite, the **DarkGridV1** suite does not vary  $h_{100}, n_s, \Omega_b$ , or the neutrino mass. This implies that the distribution of these parameters when the two sets of simulations are combined is narrower than that arising from the Gower Street suite alone. We verified in Appendix D that this has a negligible impact on our posteriors.

Last, for testing purposes only, we also use a subset of the simulations from the **CosmoGridV1** suite [45]. From this suite we chose a set of 100 full-sky simulations at the fiducial cosmology  $\sigma_8 = 0.84, \Omega_m = 0.26, w = -1, h = 0.6736, \Omega_b = 0.0493, n_s = 0.9649$ . Each simulation has been post-processed with a baryonification algorithm that mimics the impact of baryons at small scales. These simulations have been used only to produce the covariance matrices for the signal-to-noise estimates and to serve as mock data measurements for testing the full

end-to-end pipeline. They have not been used for the compression of the summary statistics, nor during the cosmological inference.

The map-making procedure is detailed in paper I. In brief, for each simulation we generate noisy mock shear maps following:

$$\gamma(p) = \frac{\sum_s \bar{n}(s)[1 + b \delta(p, s)](1 + m)[\gamma(p, s) + \gamma_{\text{IA}}(p, s)]}{\sum_s \bar{n}(s)[1 + b \delta(p, s)]} + \left( \frac{\sum_s \bar{n}(s)}{\sum_s \bar{n}(s)[1 + b \delta(p, s)]} \right)^{1/2} F(p) \frac{\sum_g w_g e_g}{\sum_g w_g}. \quad (1)$$

Here  $p$  is a simulation pixel;  $s$  is a thin redshift shell;  $\gamma(p, s)$  is the noiseless shear from the shear simulation;  $\bar{n}(s)$  is the galaxy count across the whole footprint [60];  $m$  is the multiplicative shear bias that models shear measurement uncertainties [54];  $\gamma_{\text{IA}}(p, s)$  is the intrinsic alignment contribution to each pixel;  $\delta(p, s)$  is the matter overdensity in the shear simulation;  $b$  is the galaxy-matter bias of the weak lensing sample (fixed to unity);  $w_g$  and  $e_g$  are the DES Y3 galaxy weights and ellipticities for the galaxies  $g$  in pixel  $p$  (ellipticities have been randomly rotated to erase the cosmological signal of the catalogue); and  $F(p)$  is a near-unity scale factor introduced to avoid double-counting source clustering effects (see paper I for more details). The latter reads

$$F(p) = A \sqrt{1 - B\sigma_e^2(p)}, \quad (2)$$

where  $A$  and  $B$  are constants with values  $A = [0.97, 0.985, 0.990, 0.995]$  and  $B = [0.1, 0.05, 0.035, 0.035]$  (one for each tomographic bin), and  $\sigma_e^2(p)$  is the variance of the pixel noise. The intrinsic alignment term  $\gamma_{\text{IA}}(p, s)$  is:

$$\gamma_{\text{IA}}(p, s) = A_{\text{IA}} \left( \frac{1+z}{1+z_0} \right)^{\eta_{\text{IA}}} \frac{c_1 \rho_{m,0}}{D(z)} s(p, s), \quad (3)$$

with  $z_0 = 0.62$ ,  $c_1 \rho_{m,0} = 0.0134$  (Bridle & King 15),  $D(z)$  the linear growth factor, and  $s(p, s)$  the shear tidal field. We obtain  $s(p, s)$  directly from the density field  $\delta(p, s)$  by applying the (inverse) Kaiser-Squires algorithm.

This procedure is repeated for each of the four tomographic bins of the DES Y3 source catalogue. The noisy shear maps are then converted to noisy weak lensing mass maps using the same algorithm as for the data. The mock-making procedure has several free parameters: four multiplicative shear biases  $m_i$  (one for each tomographic bin), four redshift distributions  $n_i(z)$  (one for each tomographic bin), and parameters  $A_{\text{IA}}$  and  $\eta_{\text{IA}}$  controlling the amplitude and the redshift evolution of intrinsic alignment. For each DES Y3 mock catalogue we draw at random one of these parameters from their prior (see Table I); for the redshift distributions, we draw from one of the multiple realisations provided by [60], which encompass the uncertainties in the redshift calibration of the DES Y3  $n(z)$ .

### III. SUMMARY STATISTICS

The summary statistics considered in this work are: 1) second and third moments; 2) wavelet phase harmonics; 3) the scattering transform. The summary statistics are applied to ‘smoothed’ variants of the weak lensing maps, with the choice of smoothing varying according to the specific statistic employed: moments use top hat filters, while wavelet phase harmonics and the scattering transform use wavelet filters [20, 55, 82]. We provide below a brief description of these statistics; for more details see paper I.

#### A. Second and Third Moments

Second moments of weak lensing mass maps are a Gaussian statistic, while third moments reflect the skewness of the field [17, 29, 31, 64, 65, 83–85]. Briefly, we first smooth the maps using top-hat filters, and we consider eight smoothing scales  $\theta_0$  equally (logarithmically) spaced from 8.2 to 221 arcmin. The second and third moments estimators are:

$$\langle \hat{\kappa}_{\theta_0}^2 \rangle(i, j) = \text{Avg}_p \left( \kappa_{\theta_0, p}^i \kappa_{\theta_0, p}^j \right) \quad (4)$$

$$\langle \hat{\kappa}_{\theta_0}^3 \rangle(i, j, k) = \text{Avg}_p \left( \kappa_{\theta_0, p}^i \kappa_{\theta_0, p}^j \kappa_{\theta_0, p}^k \right), \quad (5)$$

where  $\kappa_{\theta_0, p}^i$  is the smoothed lensing mass map of tomographic bin  $i$  ( $i, j, k$  refer to different tomographic bins), and the average is over all pixels  $p$  on the full sky. In the case of third moments, we subtracted noise-signal third moments of the form  $\langle \hat{\kappa}_{\theta_0, \text{obs}}^i \hat{\kappa}_{\theta_0, \text{N}}^j \hat{\kappa}_{\theta_0, \text{N}}^k \rangle$ ; this approach was chosen because it improved the compression process and minimized the influence of source clustering on our summary statistics.

#### B. Wavelet Phase Harmonics

Wavelet phase harmonics (WPH) correspond to the second moments of smoothed weak lensing mass maps that have undergone a nonlinear transformation [4, 20, 55, 82]. A directional, multi-scale wavelet transform is used to smooth the maps. Let us consider a smoothed map  $\kappa_{n, \ell}^i(\vec{\theta})$ , where  $n$  specifies the size of the filter (roughly equivalent to  $2^{n+1}$  pixels), and  $\ell$  the orientation. The non-linear operation on the smoothed map, which enables us to capture inter-scale interactions and access the non-Gaussian characteristics of the field through second moments, is defined as:

$$\text{PH}(r e^{i\theta}, q) \equiv r e^{iq\theta}, \quad (6)$$

where  $r$  is the modulus of the field and  $\theta$  its phase. We consider only  $q = 0$  or  $q = 1$  in this work, corresponding to taking the modulus or leaving the field unaltered,

respectively. The WPH statistics used in this work are:

$$S00(i, j, n) = \text{Avg}_p \text{Avg}_\ell \left( |\kappa_{n,\ell}^i| |\kappa_{n,\ell}^j| \right) \quad (7)$$

$$S11(i, j, n) \equiv \text{WPHG}(i, j, n) = \text{Avg}_p \text{Avg}_\ell \left( \kappa_{n,\ell}^i \kappa_{n,\ell}^j \right) \quad (8)$$

$$S01(i, j, n) = \text{Avg}_p \text{Avg}_\ell \left( |\kappa_{n,\ell}^i| \kappa_{n,\ell}^j \right) \quad (9)$$

$$C01\delta\ell0(i, j, n_1, n_2) = \text{Avg}_p \text{Avg}_\ell \left( |\kappa_{n_1,\ell}^i| |\kappa_{n_2,\ell}^j| \right) \text{ for } n_1 < n_2 \quad (10)$$

$$C01\delta\ell1(i, j, n_1, n_2) = \text{Avg}_p \text{Avg}_\ell \left( |\kappa_{n_1,\ell+1}^i| |\kappa_{n_2,\ell}^j| \right) \text{ for } n_1 < n_2. \quad (11)$$

We average over all pixels  $p$  and we also average over the three values 0, 1, 2 of the rotation index  $\ell$  (corresponding to three possible orientations of the directional wavelet). The number  $n$  varies from 0 to 5. The above summary statistics probe both Gaussian and non-Gaussian features of the field (with the exception of WPHG, which is equivalent to the power spectrum of  $\kappa$ ). To reduce the impact of source clustering, for WPH  $S01$ ,  $C01\delta\ell0$ , and  $C01\delta\ell1$  we subtract a term involving one noise-only map and the observed noisy convergence map.

### C. Scattering Transform

The scattering transform [16, 19, 56, 80, 81] conceptually resembles the wavelet phase harmonics (WPH) previously introduced. It involves smoothing the field using a directional, multi-scale wavelet transform (same as implemented in WPH), but then followed by a modulus operation. This transform-then-modulus operation is iteratively applied  $m$  times, after which we take an overall field average; the result is the scattering transform coefficient  $ST_m$ . This study considers scattering coefficients of orders  $m = 1, 2$ . Given a directional multi-scale wavelet  $\psi_{n,\ell}$  and the convergence map  $\kappa^i$  of tomographic bin  $i$ , we obtain:

$$ST_1(i, n) = \text{Avg}_p \text{Avg}_\ell \left( |\kappa^i * \psi_{n,\ell}| \right) \quad (12)$$

$$ST_2(i, n_1, n_2, \ell') = \text{Avg}_p \text{Avg}_\ell \left( ||\kappa^i * \psi_{n_1,\ell}| * \psi_{n_2,\ell'-\ell}| \right) \text{ for } n_1 \leq n_2, \quad (13)$$

where the average runs over all the pixels  $p$  and over all values of the rotation index  $\ell$  (similar to the WPH case).

### D. Measurement in data and signal-to-noise

We present various statistics in Fig. 1, both as measured in the data and as measured in a set of simulations that we did not use for our cosmological inference

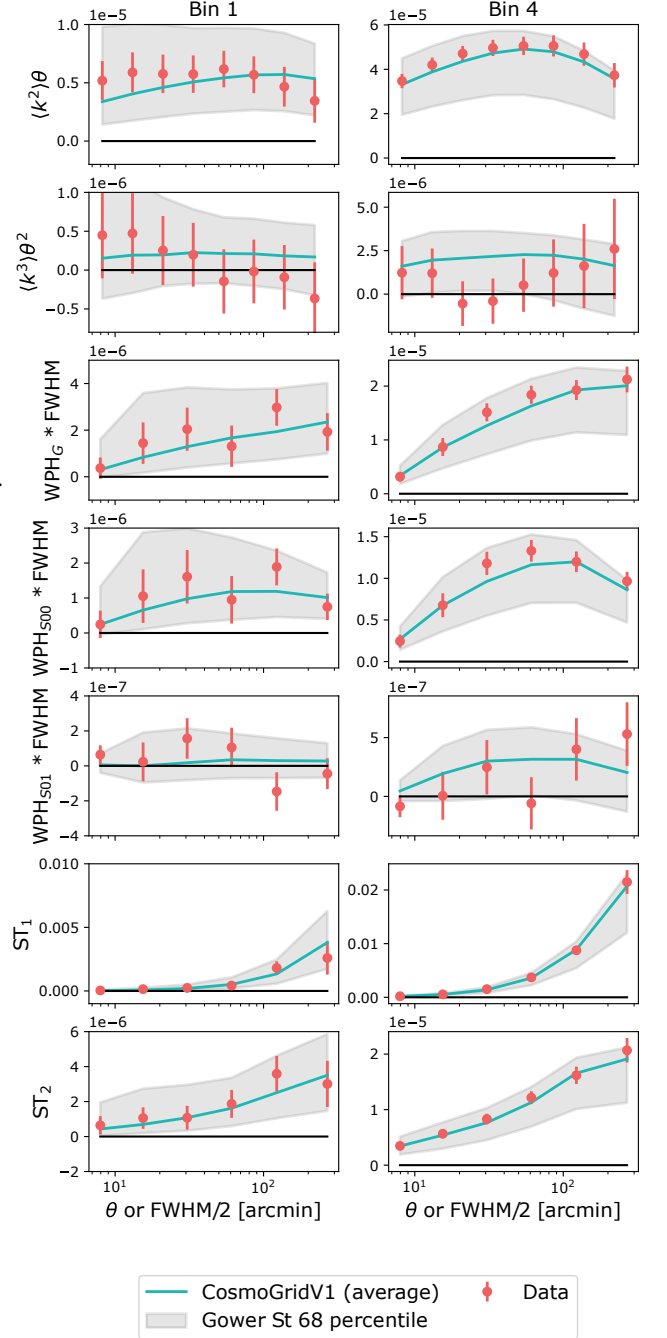


Figure 1. Some of the Gaussian and non-Gaussian statistics considered in this work, for the first (left column) and fourth (right column) tomographic bins. Red points are the statistic as measured in data, using the noisy convergence map; error bars have been estimated using 400 measurements at a fixed cosmology from the *CosmoGridV1* simulations [45]. Blue lines show the measurements in the *CosmoGridV1* simulations, averaged over the multiple simulations available at the fiducial cosmology. The blue lines only serve to guide the eye, as they are not a best fit to our data. However, the cosmological parameters of the *CosmoGridV1* simulations are not too dissimilar from our best-fit parameters, so the blue lines provide a reasonable fit. For statistics involving filters with different sizes  $j_1, j_2$ , we considered  $j_1 = j_2$ . The grey region shows for comparison the 68 % contour of the noisy measurements from the Gower Street simulations with  $|A_{IA}| < 1$  and  $|\eta_{IA}| < 1$  (which correspond to  $\sim 500$  mock measurements).

Table II. Salient properties of the summary statistics. The second column denotes whether it carries Gaussian (G) or non-Gaussian (NG) information. The third column refers to the order of the field  $\kappa$ . The fourth column is the number of components of the data vector across scales and tomographic bins. The further columns show the signal-to-noise ratio (SN) of the measurements in data.

	G/NG	Order	Data vector length	SN(Bin 1)	SN(Bin 2)	SN(Bin 3)	SN(Bin4)
2nd moments	G	2	160	3.9	7.8	16.6	13.7
3rd moments	NG	3	512	0.7	1.3	2.5	1.9
WPH S11 (WPHG)	G	2	120	3.4	7.1	15.9	13.1
WPH S00	NG	2	96	3.2	6.7	15.2	12.3
WPH S01	NG	2	480	0.5	2.5	2.1	3.2
ST1	NG	1	60	3.4	7.6	16.3	14.4
ST2	NG	1	630	3.3	7.3	15.9	13.7

pipeline (CosmoGridV1 simulations [45]). The simulated data vector is used solely to guide the eye, as the simulations are not expected to provide the best fit to the measurements in the data. The figure also shows the 68% contour spanned by the noisy measurements from the Gower Street simulations; these encompass the measurements in data very effectively. When calculating this region, we limited the analysis to an Intrinsic Alignment (IA) interval smaller than the Gower Street prior ( $|A_{IA}| < 1$  and  $|\eta_{IA}| < 1$ ) to simplify visualisation. Without this restriction, the grey region in the first bin would be significantly larger, as it can be highly impacted by large IA values. We also report the measured signal-to-noise ratio in Table II; these ratios are very similar to the ones reported in paper I for simulated measurements.

### E. Data Compression

Data compression increases the efficiency of density estimation by reducing the data vector’s dimensionality. In this work we use neural compression and we match the dimension of the compressed summary statistics to that of the parameters of interest  $\theta$ . Specifically, for a summary statistic  $\mathbf{d}$ , we compress it as  $\mathbf{t} = F_\phi(\mathbf{d})$ , modelling the compression function  $F_\phi$  with a neural network. We optimize  $\phi$  by minimizing a mean squared error (MSE) loss using the first half of our pseudo-independent mocks. The architecture used for the network and the number of parameters are summarised in Table 3 in paper I.

Given that we are principally focused on constraining the parameters  $\Omega_m$ ,  $S_8$ ,  $w$ , and  $A_{IA}$ , we target these parameters exclusively for individual compression and refrain from compressing the data vectors for any other parameters. We compress each summary statistic separately and then amalgamate their compressed forms during the likelihood inference stage. Specifically, we compress the second moments, third moments, WPHG, WPH S00, ST<sub>1</sub>, and ST<sub>2</sub> on an individual basis, whereas for optimisation purposes we compress WPH S01 and WPH C01 jointly.

## IV. SIMULATION-BASED INFERENCE AND POSTERIOR ESTIMATION

This work relies on simulation-based inference (also known as likelihood-free inference) to ensure a reliable inference of the cosmological parameters. More details are given in paper I; in brief, in simulation-based inference the likelihood is not assumed to have a closed form, but rather is estimated from mock noisy realisations of the summary statistics. We use an ensemble of neural density estimators (NDEs) – Gaussian Mixture Density Networks (MDNs; [12]) and Masked Autoregressive Flows (MAF; [61]) – to estimate the conditional distribution  $p(d|\theta)$ . NDEs approximate this distribution with estimates  $q(d|\theta; \phi')$  where the network parameters  $\phi'$  are optimized by minimizing a loss function.

We use the package pyDELFI [5] both for density estimation and for neural network training. For the training phase, we input the compressed statistics from half of our simulations, specifically those not used in the compression step. Density estimation is performed using the parameter set  $\theta = [\Omega_m, S_8, w, A_{IA}]$  and the corresponding compressed data vectors. This approach implies that the remaining parameters, which vary in the mock productions, are effectively marginalized over [43]. The process of marginalization respects the prior distributions that were applied to sample these parameters during the mock generations, as detailed in Table I.

The final density estimation combines the different ensemble estimates (MDNs and MAFs) weighted by training-derived losses. With the estimated target distribution  $p(d|\theta)$ , we compute the likelihood at observed data  $d = d_{obs}$ . The final posteriors are derived via Markov chain Monte Carlo (MCMC) sampling of the likelihood, while considering the priors mentioned in Table I for the parameters  $\theta = [\Omega_m, S_8, w, A_{IA}]$ . For the  $\Lambda$ CDM analysis, we simply run the MCMC sampling fixing  $w = -1$ . This MCMC sampling is carried out using the publicly available software package EMCEE [28], an ensemble sampler with affine-invariant properties designed for MCMC.

## V. PRE-UNBLINDING TESTS

Blinding procedures are commonly used in weak lensing analyses to prevent observer biases. Researchers intentionally conceal from themselves specific details of the data or analysis pipeline until late in the analysis; this promotes objectivity and reduces the influence of expectations on the results. The main DES Y3 weak lensing analysis [6, 76] uses a two-stage blinding scheme. First, the weak lensing sample was blinded by means of a multiplicative factor [34]; second, the summary statistics under study were manipulated according to a transformation introduced by [59]. This transformation was specifically designed to induce a shift in the posterior distributions, while ensuring that the measured data vectors could still be subjected to systematic tests and analysis. At the time of writing, the shape catalogue was already published and unblinded; therefore the first level of blinding was not applied in this paper. Furthermore, our pipeline does not accommodate the second blinding stage, as that stage relies on a parameter-based model of noise-free summary statistics, which we lack. Given that the main DES Y3 cosmological findings are already public, we opted not to blind our catalogue or data vectors. Instead, we refrained from examining the data posteriors until after our pipeline validation and null tests on the data vectors were successfully completed. In addition to the tests conducted at the catalogue and map level, as described in [34] and [42], we performed the following tests, detailed in paper I:

- **Validation of the Gower Street simulations:** We tested that the power spectra of the noiseless, full-sky convergence of the Gower Street simulations were in an approximately one percent agreement with theory predictions; moreover, we tested that the noise properties of our mocks (i.e. second, third, and fourth moments of the noise, as well as its cumulant distribution function) were able to describe the noise properties of our data.
- **Validation of the posteriors:** We checked, using an empirical coverage test, that the size of the posteriors estimated by our NDEs was not misestimated.
- **End-to-end validation of the pipeline on simulations:** We tested, using an independent set of simulations, that the pipeline was able to recover true cosmology for the summary statistics considered in this work.
- **Tests on baryonic contamination:** We conducted tests using a subset of simulations that have been post-processed to incorporate baryonic feedback via the baryonic correction model [8, 74]. Our objective was to verify that our statistics and the scales utilized in this study were not significantly impacted by potential baryonic contamination, which the simulations in our pipeline do not model. Specifically, we confirmed that the peak

of the marginalized two-dimensional posterior distribution of  $\Omega_m$  and  $S_8$ , when analyzing baryon-contaminated data, falls within  $0.3\sigma$  of the peak obtained from clean data.

- **Tests on additive biases due to PSF modelling errors:** We checked that additive biases due to PSF modelling errors were negligible at the data vector level, i.e. that if neglected they would not bias our cosmological analysis. This test is similar to that performed in the DES Y3 cosmic shear analysis [6].
- **Tests on possible modelling errors of the source clustering effect:** We checked that potential errors in our model for the source clustering effects were negligible, i.e. that they would not bias our cosmological analysis.

These tests were performed in paper I for a  $w$ CDM analysis. Appendix A repeats some of these for a  $\Lambda$ CDM analysis, and also adds two sets of tests:

- **Test on the impact of redshift uncertainties and shear biases:** We show that we can recover the true cosmology of the simulations even if the parameters describing redshift uncertainties and shear bias calibrations were  $2\sigma$  off their mean values.
- **Sensitivity to the details of the N-body simulations:** We show that our posteriors are not sensitive to the details of the simulations used (box size, number of particles, redshift resolution).

Finally, before examining the posteriors we performed three additional tests:

- **B-mode ‘null-test’:** Weak lensing does not produce B-modes at first order, making such modes useful for identifying systematics. For masked data, however, the map-making procedure can cause small B-modes due to E-modes leakage at large scales. We perform in Appendix B a comparison between data and simulated B-modes; this suggests no anomaly in our data measurements.
- **Goodness-of-fit:** We assess the goodness-of-fit  $p$ -value for the compressed data vectors, ensuring they do exceed 1 percent ( $p$ -value  $> 0.01$ , see Section VI, Tables IV, III). We used the following procedure to measure  $p$ -values. Let  $\mathcal{L}(\mathbf{x})$  denote the likelihood of a data vector  $\mathbf{x}$ . For each summary statistic (or combination of statistics), we trained the Neural Density Estimators on all (compressed) simulated data vectors except one, denoted  $\mathbf{x}_i$ , selected at random. We then estimated the minimum log likelihood<sup>2</sup> from the learned likelihood surface for the simulated data vector excluded from the

<sup>2</sup> The routine `scipy.optimize.differential_evolution` was used.



training; we denote this minimum log likelihood by

$$\min\text{LL}_i \equiv \min \log \mathcal{L}(\mathbf{x}_i). \quad (14)$$

This process was repeated for a thousand data vectors. Next, we estimated the minimum log likelihood for our data:

$$\min\text{LL}_{\text{data}} = \min \log \mathcal{L}(\mathbf{d}). \quad (15)$$

Subsequently, we computed the probability-to-exceed  $p$  by counting the fraction of simulated data vectors that had a minimum log likelihood greater than that of the observed data:

$$p = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\min\text{LL}_i > \min\text{LL}_{\text{data}}), \quad (16)$$

where  $\mathbb{1}$  is the indicator function, which is 1 if its argument is true and 0 otherwise. We note that this goodness-of-fit test is valid only for compressed data vectors; a satisfactory  $p$ -value for compressed statistics does not necessarily imply a similar outcome for uncompressed statistics. Unfortunately, we are unable to test the uncompressed statistics, as learning the likelihood of the uncompressed measurements is not feasible with the limited simulations available to us. However, we note that the uncompressed data are effectively encompassed by the uncompressed mock measurements from the Gower Street simulations, as depicted in Fig. 1. This provides evidence that the uncompressed data vector does not exhibit features that are unaccounted for in our mock simulations.

- **Stability of the posterior:** We visually checked that the four NDEs were delivering consistent posteriors when considered individually (Appendix ??). This was done by blinding the cosmological parameter axes.
- **Internal consistency of summary statistics:** We assess the internal consistency of the various summary statistics employed. To do this, we use the Posterior Predictive Distribution (PPD) methodology, developed by [24] and implemented in the main DES Y3 analysis. This methodology is designed to calculate a calibrated probability-to-exceed  $p$ -value. For consistency tests between two types of summary statistics, we generate realisations of one type of summary statistic (e.g. second moments) under the assumption that the other type (e.g. third moments) is the one measured in data. These realisations are obtained using parameters drawn from the posterior combining both types of statistics. In our case, we easily obtained these realisations from the joint likelihood, which we learned using the noisy (compressed) measurements from the Gower Street simulations. These realisations are then compared to actual observations using a distance metric (in our case, the log

likelihood) in the compressed data space; this metric is then used to determine the  $p$ -value. More details on the exact implementation are given in [24]. We report in Table V the PPD  $p$ -values obtained comparing second moments against third moments, ST, and WPH (and vice-versa); all the values are well above the  $p = 0.01$  threshold. Note that the PPD calculation is not symmetric in its two arguments.

Only once all these tests were passed did we look at the unblinded posteriors of our analysis.

## VI. RESULTS

We present the cosmological constraints obtained from our analysis of the DES Y3 data, assuming in turn the  $\Lambda$ CDM and  $w$ CDM models. We focus on the parameters used to learn the likelihood, which are the most constrained ones:  $\Omega_m$ ,  $S_8$ ,  $A_{1A}$ , and (for the  $w$ CDM scenario)  $w$ . The other cosmological parameters and nuisance parameters varied when creating the mocks are automatically marginalized over in our formalism. When reporting constraints, we also report the Figure of Merit (FoM) associated with  $S_8$  and  $\Omega_m$ , defined by

$$\text{FoM}(S_8, \Omega_m) = (\det(C_{S_8, \Omega_m}))^{-0.5}. \quad (17)$$

where  $C_{S_8, \Omega_m}$  is the covariance matrix of the parameters (estimated from their posterior).

### A. $\Lambda$ CDM results

This subsection presents the  $\Lambda$ CDM results. The left plot in Fig. 2 displays the posterior distributions for  $S_8$ ,  $\Omega_m$ , and  $\sigma_8$  from a combination of different summary statistics; the constraints are also detailed in Table III, together with  $p$ -values and the FoMs. Constraints for the individual statistics are shown in Appendix E. As confirmed by our PPD consistency tests (Table V), all the posteriors appear to be consistent with each other, with their constraints largely overlapping. The marginalised mean values of  $S_8$ ,  $\Omega_m$ , and  $\sigma_8$  for the combination of all the summary statistics (grey line in plot), along with the 68 percent credible intervals, are:

$$\begin{aligned} S_8 &= 0.794 \pm 0.017 \\ \Omega_m &= 0.259 \pm 0.025 \\ \sigma_8 &= 0.857 \pm 0.042. \end{aligned}$$

The constraints obtained by combining all the summary statistics are significantly tighter than the constraints from second moments only: by  $\sim 15$  percent for  $S_8$ ,  $\sim 25$  percent for  $\sigma_8$  and  $\Omega_m$ , and  $\sim 70$  percent for the FoM. Importantly, constraints significantly improve not only  $S_8$  but also  $\Omega_m$ . When combining non-Gaussian statistics

Table III. Constraints on various parameters for different summary statistics and their combinations, for the  $\Lambda$ CDM model. For each parameter we report the mean and 68 percent credible interval. The last two columns report the FoM and the  $p$ -values.

$\Lambda$ CDM Summary Statistic(s)	$S_8$	$\sigma_8$	$\Omega_m$	$A_{IA}$	FoM( $S_8, \Omega_m$ )	$p$ -value
2nd moments	$0.796^{+0.021}_{-0.020}$	$0.849^{+0.055}_{-0.056}$	$0.267^{+0.033}_{-0.033}$	$0.45^{+0.48}_{-0.42}$	1303	0.39
2nd+3rd moments	$0.801^{+0.021}_{-0.021}$	$0.847^{+0.044}_{-0.044}$	$0.270^{+0.028}_{-0.029}$	$0.51^{+0.53}_{-0.50}$	1398	0.79
2nd moments+ST	$0.791^{+0.019}_{-0.018}$	$0.846^{+0.044}_{-0.044}$	$0.264^{+0.027}_{-0.028}$	$0.34^{+0.35}_{-0.33}$	1770	0.41
2nd moments+WPH	$0.801^{+0.020}_{-0.020}$	$0.866^{+0.044}_{-0.045}$	$0.259^{+0.027}_{-0.026}$	$0.44^{+0.45}_{-0.43}$	1817	0.75
2nd moments+ST+WPH	$0.791^{+0.018}_{-0.018}$	$0.855^{+0.043}_{-0.044}$	$0.259^{+0.027}_{-0.027}$	$0.36^{+0.37}_{-0.36}$	2069	0.60
2nd+3rd moments+ST+WPH	$0.794^{+0.017}_{-0.017}$	$0.857^{+0.042}_{-0.042}$	$0.259^{+0.025}_{-0.025}$	$0.32^{+0.33}_{-0.31}$	2234	0.71
DES Y3 cosmic shear	$0.788^{+0.016}_{-0.019}$	$0.810^{+0.061}_{-0.071}$	$0.289^{+0.044}_{-0.039}$	$0.33^{+0.28}_{-0.27}$	1402	
DES Y3 2nd+3rd moments (G22)	$0.787^{+0.021}_{-0.016}$	$0.819^{+0.038}_{-0.028}$	$0.279^{+0.033}_{-0.029}$	$0.40^{+0.61}_{-0.38}$	1747	
DES Y3 Cl+Peaks (Z22)	$0.797^{+0.014}_{-0.014}$	$0.849^{+0.100}_{-0.117}$	$0.276^{+0.077}_{-0.064}$	$-0.04^{+0.19}_{-0.19}$	1031	

Table IV. Constraints on various parameters for different summary statistics and their combinations, for the  $w$ CDM model. For each parameter we report the mean and the 68 percent credible interval (except that for  $w$  we report the mean and 68 percent upper limit). The last two columns report the FoM and the  $p$ -values.

$w$ CDM Summary Statistic(s)	$S_8$	$\sigma_8$	$\Omega_m$	$w$	$A_{IA}$	FoM( $S_8, \Omega_m$ )	$p$ -value
2nd moments	$0.826^{+0.028}_{-0.028}$	$0.870^{+0.056}_{-0.057}$	$0.274^{+0.037}_{-0.037}$	$< -0.62$	$0.31^{+0.37}_{-0.29}$	899	0.66
2nd+3rd moments	$0.825^{+0.026}_{-0.027}$	$0.858^{+0.047}_{-0.047}$	$0.280^{+0.033}_{-0.033}$	$< -0.69$	$0.58^{+0.39}_{-0.39}$	1121	0.73
2nd moments+ST	$0.814^{+0.024}_{-0.024}$	$0.847^{+0.046}_{-0.046}$	$0.279^{+0.032}_{-0.033}$	$< -0.68$	$0.38^{+0.30}_{-0.29}$	1274	0.30
2nd moments+WPH	$0.830^{+0.024}_{-0.024}$	$0.862^{+0.043}_{-0.043}$	$0.281^{+0.032}_{-0.032}$	$< -0.65$	$0.35^{+0.31}_{-0.30}$	1435	0.90
2nd moments+ST+WPH	$0.819^{+0.021}_{-0.021}$	$0.855^{+0.045}_{-0.045}$	$0.277^{+0.032}_{-0.032}$	$< -0.70$	$0.28^{+0.25}_{-0.25}$	1605	0.58
2nd+3rd moments+ST+WPH	$0.817^{+0.021}_{-0.021}$	$0.861^{+0.041}_{-0.041}$	$0.273^{+0.029}_{-0.029}$	$< -0.72$	$0.37^{+0.30}_{-0.29}$	1725	0.67
DES Y3 cosmic shear	$0.813^{+0.032}_{-0.022}$	$0.816^{+0.082}_{-0.077}$	$0.303^{+0.040}_{-0.051}$	$< -0.77$	$0.34^{+0.28}_{-0.25}$	901	

$\Lambda$ CDM Data Splits	$p$ -values
2nd moments <i>vs.</i> 3rd moments	0.50
3rd moments <i>vs.</i> 2nd moments	0.20
2nd moments <i>vs.</i> ST	0.58
ST <i>vs.</i> 2nd moments	0.28
2nd moments <i>vs.</i> WPH	0.16
WPH <i>vs.</i> 2nd moments	0.11
$w$ CDM Data Splits	$p$ -values
2nd moments <i>vs.</i> 3rd moments	0.47
3rd moments <i>vs.</i> 2nd moments	0.36
2nd moments <i>vs.</i> ST	0.45
ST <i>vs.</i> 2nd moments	0.33
2nd moments <i>vs.</i> WPH	0.14
WPH <i>vs.</i> 2nd moments	0.13

Table V. Summary of internal consistency test  $p$ -values. All internal consistency tests pass the pre-defined (arbitrary) threshold of 0.01.

with second moments, WPH provides the largest gain, followed by ST and 3rd moments; the combination of all the summary statistics provide a further enhancement in the constraints, confirming the findings of paper I. The gain in constraining power comes both from the additional information probed by non-Gaussian statistics and from the degeneracy breaking in the  $\Omega_m$ - $\sigma_8$  plane (as non-Gaussian statistics are characterized by a slightly different tilt in that plane compared to second moments).

Fig. 3 shows the constraints on the intrinsic alignment parameter  $A_{IA}$ , for second moments and the combination of all summary statistics. The constraints improve by up to approximately 35 percent when non-Gaussian statistics are included, underscoring the ability of beyond-Gaussian statistics to enhance non-cosmological parameters as well. Our measurements indicate a preference for a small positive (i.e. non-zero) intrinsic alignment amplitude.

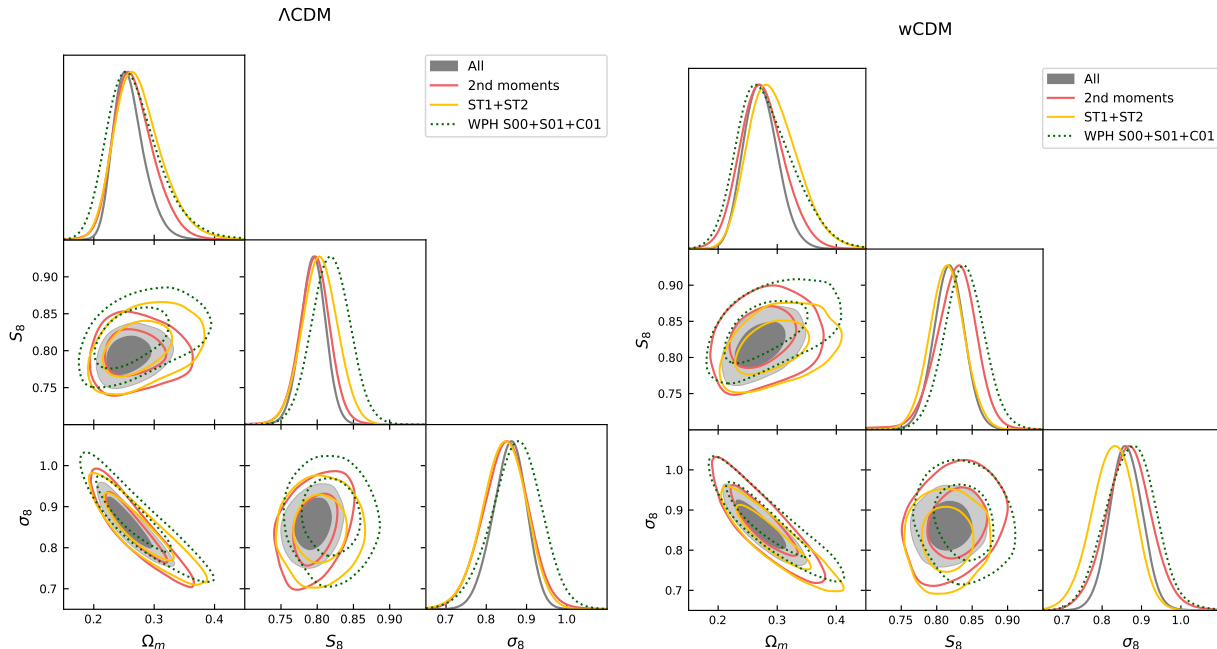


Figure 2. Posterior distributions of the cosmological parameters  $\Omega_m$ ,  $S_8$ , and  $\sigma_8$ , for different summary statistics and their combinations (“All”), as measured in our data ( $\Lambda$ CDM on the left,  $w$ CDM on the right). The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions.

### B. $w$ CDM results

This subsection analyzes the  $w$ CDM results. The right plot in Fig. 2 displays the posterior distributions for  $S_8$ ,  $\Omega_m$ , and  $\sigma_8$  for a combination of summary statistics (the same as shown for the  $\Lambda$ CDM case); constraints are also detailed in Table IV, along with  $p$ -values and the Figures of Merit (FoMs). Fig. 4 shows the posterior for  $S_8$  and  $w$ . The results are in spirit very similar to what we obtained for the  $\Lambda$ CDM case; posteriors derived from different summary statistics largely overlap, and the combination of these statistics significantly improves constraints compared to those from second moments only. The marginalised mean values of  $S_8$ ,  $\Omega_m$ , and  $\sigma_8$  for the combination of all the summary statistics (grey line in plot), along with the 68 percent credible intervals, are:

$$\begin{aligned} S_8 &= 0.817 \pm 0.021 \\ \Omega_m &= 0.273 \pm 0.029 \\ \sigma_8 &= 0.861 \pm 0.041. \end{aligned}$$

We observe an improvement in the constraints over second moments of approximately 25 percent for  $S_8$  and  $\sigma_8$ , approximately 20 percent for  $\Omega_m$ , and about 90 percent for the Figure of Merit (FoM) when all the statistics are considered. Interestingly, we obtain slightly higher values for  $S_8$  in the  $w$ CDM scenario compared to the  $\Lambda$ CDM case. This is a consequence of opening the parameter space along  $w$  in a non-symmetric way ( $w > -1$ );

since the  $S_8$  and  $w$  posteriors show a slight degeneracy (see Fig. 4), this pushes  $S_8$  upwards compared to the case when  $w$  is fixed at  $-1$ . Concerning the constraints on the parameter  $w$ , all the probes are consistent with each other. The constraints get smaller as we include more summary statistics, and they push towards our prior boundary at  $w = -1$ , consistent with a  $\Lambda$ CDM scenario. Last, the constraints on the intrinsic alignment amplitude  $A_{IA}$  are completely consistent with the ones obtained in the  $\Lambda$ CDM case, as shown in Fig. 3.

### C. Comparison with other DES analyses

We discuss how the parameter constraints obtained from this work compare to those obtained by other cosmological analyses using DES Y3 weak lensing data. Several analyses incorporating Gaussian and non-Gaussian statistics have been conducted with DES Y3 data. Although these studies use identical data and calibration, direct comparison of their outcomes is complex. This complexity arises from varying analysis and modelling choices adopted in each study, as these can significantly impact the results. To facilitate a more accurate comparison, we attempted to replicate the analyses using similar analysis and modelling choices as in our study, wherever feasible. Below, we provide an overview of each study we are comparing. Fig. 5 offers a visual comparison of the constraints on  $S_8$ ,  $\sigma_8$ , and  $\Omega_m$  for both the  $\Lambda$ CDM and  $w$ CDM models, and the corresponding constraint values and Figures of Merit (FoMs) are detailed

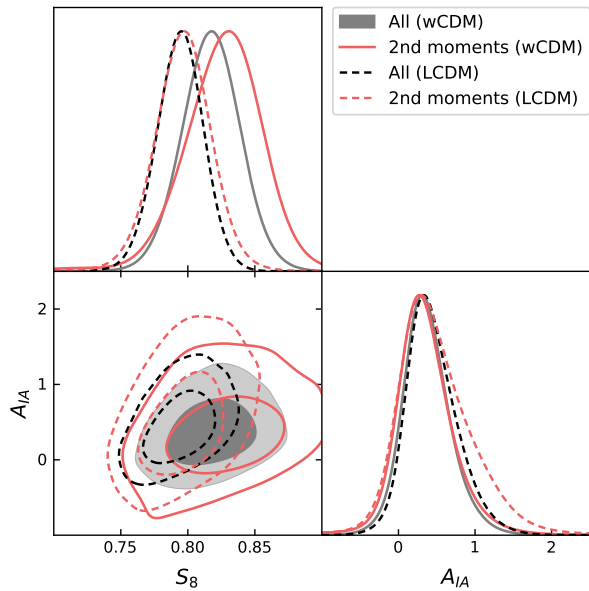


Figure 3. Posterior distributions of the cosmological parameters  $S_8$  and  $A_{IA}$ , for second moments and for the combination of all the summary statistics (“All”), to better highlight the contribution of non-Gaussian statistics to the overall constraints. The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions.

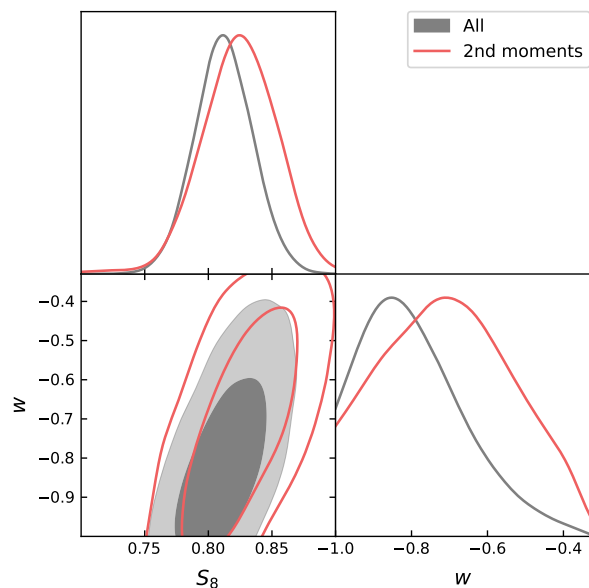


Figure 4. Posterior distributions of the cosmological parameters  $S_8$  and  $w$ , for second moments and for the combination of all the summary statistics (“All”), to better highlight the contribution of non-Gaussian statistics to the overall constraints. The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions.

in Tables III and IV.

*DES Y3 cosmic shear.* The first study to which we compare our results are the DES Y3 cosmic shear analyses [6, 76] for both  $\Lambda$ CDM and  $w$ CDM models. We repeated these analyses trying to match the same analysis choices used in our study. Notably, we excluded the shear ratio likelihood [73], since it is not included in our pipeline. Shear ratios, which are galaxy-galaxy lensing measurements at small scales, primarily constrain intrinsic alignment (IA) and redshift parameters. Since IA amplitude is correlated with  $S_8$ , incorporating shear ratios significantly enhances  $S_8$  constraints. We also adopted the same priors for parameters  $n_s$ ,  $\Omega_b$ ,  $h_{100}$ , and neutrino mass used in this work. For redshift uncertainties, we used the HYPERRANK method [21], which was not the standard method in [6, 76], but which is employed in our current analysis. We also used as intrinsic alignment model the Non-Linear Alignment (NLA) model, augmented with a clustering term with a fixed galaxy-matter bias of one. We note, however, that the clustering term included in the IA model of the cosmic shear analysis is estimated using tree-level perturbation theory, whereas our implementation directly uses the clustering of the simulation, which should be more accurate. Another difference with our pipeline is that it automatically excludes information on large multipoles ( $\ell > 1000$ ) by imposing a hard cut during the map making process. This restriction is not present in the cosmic shear analysis, which can theoretically probe smaller scales. Both analyses, however, have undergone the same process of scale cuts, where small scales were excluded if they were potentially affected by unmodelled baryonic effects. Last, the full impact of source clustering is not accounted for in the cosmic shear analysis, although its impact should be negligible [33].

Fig. 5 and Tables III and IV show the remarkable agreement between the DES cosmic shear analysis and our work. The cosmic shear Figure of Merit (FoM) is very similar to that from our second moments; this is expected, as they are both Gaussian statistics applied to the same dataset with similar analysis choices. The FoM from our work, when all summary statistics are combined, greatly improves over the cosmic shear one, as expected. The constraints on intrinsic alignment and  $w$  are also compatible. Interestingly, despite the FoM from the combination of all statistics being almost double that of the cosmic shear analysis, we note that not all parameter constraints are necessarily improved (e.g.  $A_{IA}$  in the  $w$ CDM case or  $S_8$  in the  $\Lambda$ CDM case). This might be due to the limited access to small-scale information at  $\ell > 1000$  not probed by our work, a non-optimal compression of our statistics along the  $S_8$  and  $A_{IA}$  directions, or simply due to a slightly different degeneracy direction in the  $\sigma_8$ – $\Omega_m$  plane, a consequence of different constraining power. Despite this, the significant gain in the FoM over the fiducial DES cosmic shear analysis, and the compatibility of the constraints obtained using a completely different and independent pipeline, strongly highlight the importance of a simulation-based analysis

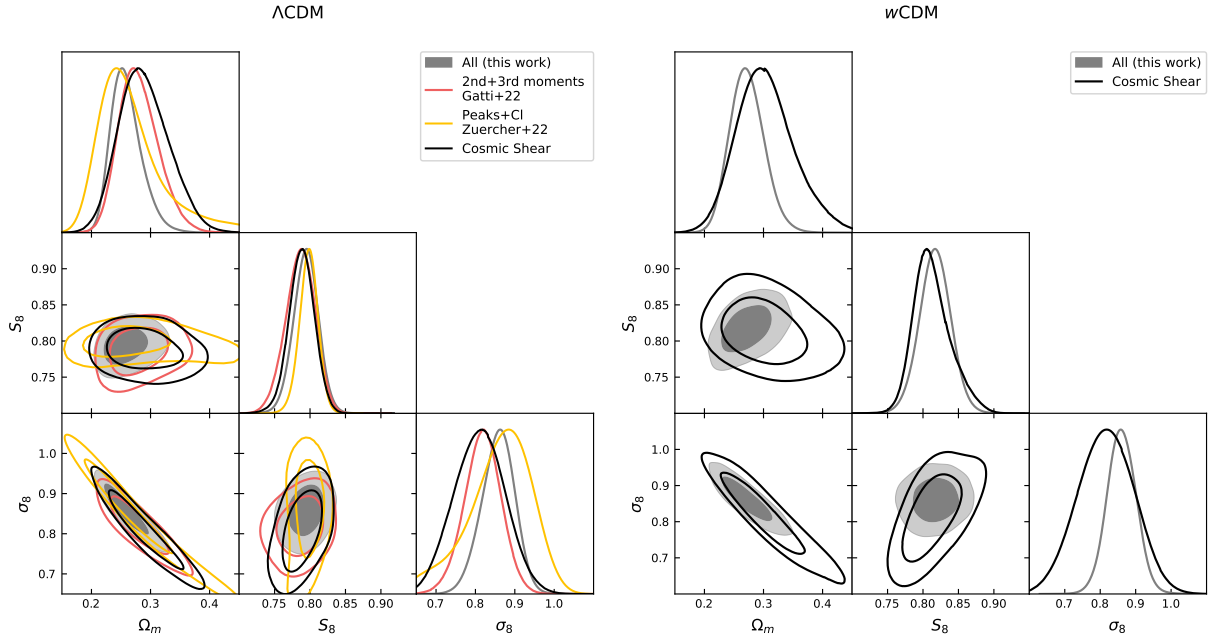


Figure 5. Posterior distributions of the cosmological parameters  $\Omega_m$ ,  $S_8$ , and  $\sigma_8$ , as measured in data. We compare the results of this work to other analyses using DES data. The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions.

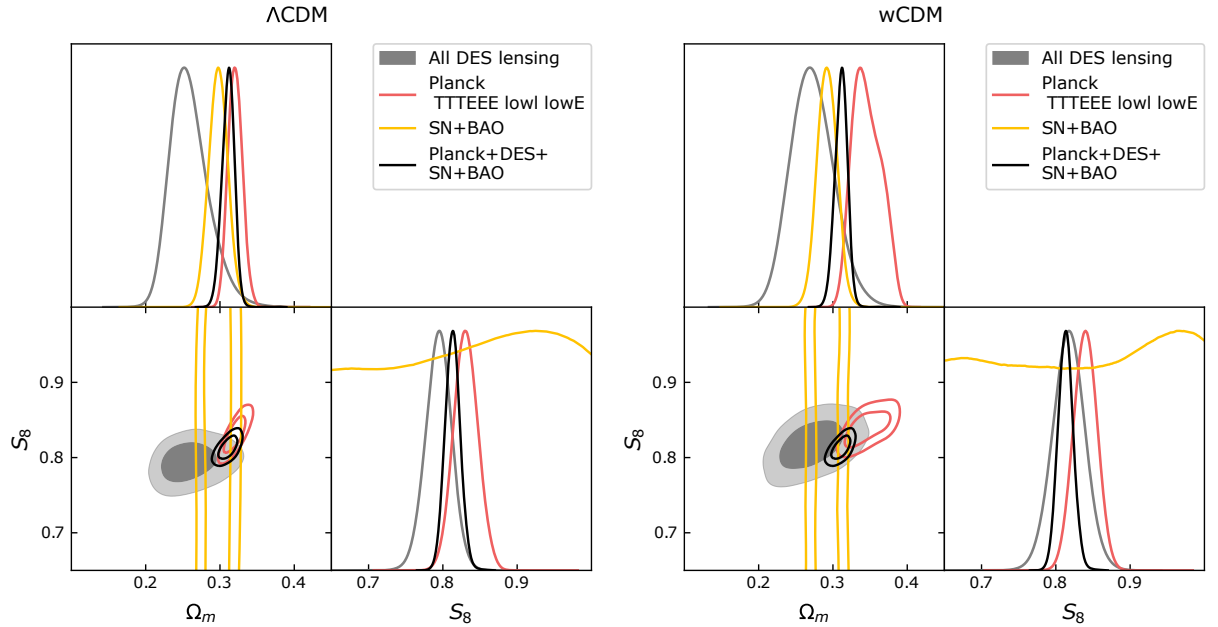


Figure 6. DES Y3 and external data constraints from low and high redshift probes. The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions. We find no evidence for statistical inconsistency between the results of this work and external data set.

using non-Gaussian statistics.

*DES Y3 second and third moments.* The second study to which we compare our results is the  $\Lambda$ CDM analysis of second and third moments from Gatti et al. [31]. This analysis uses the third moments as non-

Gaussian statistics, measured in the same way as we did in this work. However, the modelling approach was analytical, in contrast to our pipeline. Additionally, a MOPED compression [37] was used, in contrast to the neural network compression used in the current study.

As with the case of cosmic shear, we endeavoured to align the analysis choices with our pipeline. In contrast to the fiducial analysis choices in [31], we reran the second and third moments analysis without including shear ratios; furthermore, we matched the priors on the parameters  $n_s$ ,  $\Omega_b$ , and  $h_{100}$ . However, the [31] analysis did not vary the neutrino mass, as neutrinos were not modelled; this was an analysis choice we could not alter. Additionally, source clustering effects were not modelled, although their impact on third moments was mitigated using the method explained in [33]. For the intrinsic alignment model, we employed the NLA model, but we were unable to include a galaxy-matter bias term. Lastly, as in the cosmic shear analysis, no explicit cut on multipoles  $\ell > 1000$  was made, though a scale cut intended to minimize the impact of unmodelled baryonic effects was implemented.

Results from this revised analysis of the second and third moments also overlap with the constraints from our work (see Fig. 5, and Tables III and IV). The Figure of Merit (FoM) is larger than that of our second moments, but smaller than the one from the combination of all summary statistics. The constraints are also compatible with those obtained from our implementation of the second and third moments analysis. Given the subtle differences between the two pipelines, we did not expect a perfect match; however, the match is notably good. This serves as further validation of the work done here and the analytical modelling in [31], considering that the two pipelines are completely independent and follow very different approaches.

*DES Y3 Peak counts.* The third study to which we compare our results are the  $\Lambda$ CDM constraints from the power spectrum plus peak counts analysis presented in Zuercher et al. [89]. This analysis includes peak counts as a non-Gaussian statistic and is entirely simulation-based, though it differs from our work in several aspects.

Firstly, the analysis relied solely on the `DarkGridV1` simulation suite, primarily exploring two cosmological parameters ( $S_8$  and  $\Omega_m$ ). The dependence on other cosmological parameters ( $n_s$ ,  $\Omega_b$ , and  $h_{100}$ ) was modelled through a Taylor expansion of the data vector, calibrated against a limited set of N-body simulations. Neutrinos were not varied in this analysis. To model the data vector, a three-parameter emulator was constructed for  $S_8$ ,  $\Omega_m$ , and  $A_{IA}$ ; other nuisance parameters were modelled by Taylor-expanding the data vector, an approach different from ours. A Non-Linear Alignment (NLA) intrinsic alignment (IA) model was used, without the clustering term, and source clustering effects were not included. As in our work, no shear ratio likelihood was incorporated in the original analysis. No explicit cut on multipoles  $\ell > 1000$  was made, although a scale cut was implemented to reduce the impact of unmodelled baryonic effects. A MOPED compression was used to compress the data vector, and a Gaussian likelihood was assumed. In our comparison with the peaks analysis, we retained all the fiducial choices made by [89]. We refrained from imposing the Gower Street prior on  $n_s$ ,  $\Omega_b$ , and  $h_{100}$ , as this

would have tightened the constraints and invalidated the original scale cut. For the cosmic shear and second and third moments analyses, such an adjustment was feasible as we chose to not include shear ratios, which counter-balanced the gain in constraining power from having a tight prior on  $n_s$ ,  $\Omega_b$ , and  $h_{100}$ .

The results of the peak counts analysis, as displayed in Fig. 5, are in alignment with our findings, with the posterior largely overlapping ours. Notably, this analysis delivers tighter constraints on  $S_8$  but is much less restrictive regarding  $\Omega_m$  and the overall Figure of Merit. The noticeable differences in constraining power, given the significant differences in analysis choices, are somewhat expected. This underscores the challenges in comparing results from the same dataset when different analysis methods are employed.

## VII. COMPARISON WITH EXTERNAL DATA

We compare our cosmological constraints with results from various external datasets. Although different choices could be made regarding which data to compare, we have adhered to the comparisons used in the fiducial DES cosmic shear analysis [76]. In particular we consider:

- **eBOSS:** we use spectroscopic baryon acoustic oscillation (BAO) measurements from eBOSS [3]. The BAO likelihood is assumed to be independent of DES. See [1] for a detailed description of the likelihoods included (but note that unlike [1], we do not include RSD constraints from eBOSS; this is in line with the analysis choices adopted for the DES Y3 cosmic shear analysis in [76]).
- **Pantheon Supernovae:** we consider the luminosity distances from Type Ia supernovae from the Pantheon sample [75]; this data set includes 279 type Ia supernovae from the PanSTARRS Medium Deep Survey ( $0.03 < z < 0.68$ ) and samples from SDSS, SNLS, and HST. The final Pantheon catalogue includes 1048 SNe, out to  $z = 2.26$ .
- **Planck 2018 CMB:** we use the final data release of the PLANCK Cosmic Microwave Background (CMB) survey [66] including measurements of both temperature and polarization anisotropies, in the same way it is implemented in [76].

To perform a meaningful comparison, we recomputed the posteriors using the priors on cosmological parameters used in this work (Table I) for eBOSS and supernovae data. We did not do this for the PLANCK data as the priors on  $n_s$ ,  $h$ , and  $\Omega_b$  are partially determined by PLANCK constraints. Note, though, that we are mostly interested in testing the compatibility for the parameters  $S_8$ ,  $\Omega_m$ , and  $w$ ; our analysis does not constrain the other cosmological parameters, which are completely prior dominated.

### A. Quantifying tension

To estimate the agreement or disagreement between different datasets, we use the normalizing flow Monte Carlo estimate of the probability of a parameter difference as discussed in [70]. We use the implementation of this algorithm in `tensiometer` [70, 71]. In the case of uncorrelated data sets, the probability of the parameter difference is:

$$\mathcal{P}(\Delta\theta) = \int_{V_p} \mathcal{P}_A(\theta)\mathcal{P}_B(\theta - \Delta\theta)d\theta, \quad (18)$$

where  $V_p$  is the prior support and  $\mathcal{P}_A$  and  $\mathcal{P}_B$  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:

$$\Delta = \int_{\mathcal{P}(\Delta\theta) > \mathcal{P}(0)} \mathcal{P}(\Delta\theta) d\Delta\theta, \quad (19)$$

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 convolution integral in Eq. (18) is performed with a Monte Carlo algorithm. To compute Eq. (19) we first train a normalizing flow within `tensiometer` to model the probability density of parameter shifts and then we perform a Monte Carlo integral with the trained normalizing flow. For further details, we refer the reader to [70].

In the following, our results are always reported in terms of effective number of standard deviations. When considering an event with probability  $\mathcal{P}$ , this is determined by

$$n_\sigma \equiv \sqrt{2}\text{Erf}^{-1}(\mathcal{P}), \quad (20)$$

where  $\text{Erf}^{-1}$  is the inverse error function. It represents the number of standard deviations that an event with the same probability would have if it were drawn from a Gaussian distribution. This definition does not assume a Gaussian distribution of the underlying statistics and should be interpreted as a logarithmic scale for probabilities.

#### 1. Compatibility between this analysis and external data

In Table VI we show the estimate of the tension between the different DES observables we consider, the PLANCK 2018 CMB data, and the joint BAO+SNe dataset. As we can see DES measurements are broadly in agreement with these measurements. In particular, we find that no pairwise comparison exceeds  $2.2\sigma$ . This tells us that different DES data combinations agree with geometric probes on the determination of  $\Omega_m$  and  $w$  and with the CMB on all three parameters.

$\Lambda$ CDM Summary Statistic(s)	Planck 2018 BAO+SNe	
2nd moments	1.4 $\sigma$	0.7 $\sigma$
2nd+3rd moments	1.4 $\sigma$	0.8 $\sigma$
2nd moments+ST	1.7 $\sigma$	0.9 $\sigma$
2nd moments+WPH	1.2 $\sigma$	0.8 $\sigma$
2nd moments+ST+WPH	2.0 $\sigma$	1.3 $\sigma$
2nd+3rd moments+ST+WPH	1.8 $\sigma$	1.1 $\sigma$

$w$ CDM Summary Statistic(s)	Planck 2018 BAO+SNe	
2nd moments	1.8 $\sigma$	0.9 $\sigma$
2nd+3rd moments	1.6 $\sigma$	0.7 $\sigma$
2nd moments+ST	2.0 $\sigma$	0.8 $\sigma$
2nd moments+WPH	1.8 $\sigma$	1.0 $\sigma$
2nd moments+ST+WPH	1.6 $\sigma$	1.0 $\sigma$
2nd+3rd moments+ST+WPH	2.2 $\sigma$	1.0 $\sigma$

Table VI. Compatibility of different DES observables with PLANCK and BAO + SNe results, assuming the  $\Lambda$ CDM and  $w$ CDM models.

In Table VII we show the results of estimating the discrepancy between DES probes and all other joint datasets, assuming both the  $\Lambda$ CDM and  $w$ CDM models. This reassures us that all DES data combinations can be combined with all external probes. Note that (in more than one dimension) having  $A$  and  $B$  in agreement, and having both separately in agreement with  $C$ , does not guarantee that the combination of  $A$  and  $B$  will be in agreement with  $C$ . For this reason, we explicitly check for the compatibility of all DES data combinations and joint external data.

$\Lambda$ CDM Summary Statistic(s)	Planck 2018+BAO+SNe	
2nd moments	1.1 $\sigma$	
2nd+3rd moments	1.1 $\sigma$	
2nd moments+ST	1.4 $\sigma$	
2nd moments+WPH	1.1 $\sigma$	
2nd moments+ST+WPH	1.7 $\sigma$	
2nd+3rd moments+ST+WPH	1.6 $\sigma$	

$w$ CDM Summary Statistic(s)	Planck 2018+BAO+SNe	
2nd moments	1.4 $\sigma$	
2nd+3rd moments	1.2 $\sigma$	
2nd moments+ST	1.6 $\sigma$	
2nd moments+WPH	1.5 $\sigma$	
2nd moments+ST+WPH	1.6 $\sigma$	
2nd+3rd moments+ST+WPH	1.7 $\sigma$	

Table VII. Compatibility of different DES observables with PLANCK + BAO + SNe results, assuming the  $\Lambda$ CDM and  $w$ CDM models.

## 2. Compatibility between early and late times universe

We now investigate the agreement between probes of the early and late time universe. We select the DES probe combining ‘2nd+3rd moments+ST+WPH’ since it is the most constraining among the DES summary statistics. Table VIII shows broad agreement between these two probes for both the  $\Lambda$ CDM and  $w$ CDM models.

Late (DES+SNe+BAO) vs Early (Planck 2018)	
$\Lambda$ CDM	$1.7\sigma$
$w$ CDM	$2.1\sigma$

Table VIII. Compatibility of probes of the early and late time universe within the  $\Lambda$ CDM and  $w$ CDM models. For the early Universe, we consider CMB measurements from PLANCK 2018. For the late-time Universe we use the 2nd+3rd moments+ST+WPH combination of DES probes joined with SNe and BAO measurements.

## 3. Joint constraints

Having found a good level of agreement between the posterior probability distributions from this analysis and those from external probes, we can be confident that they can be combined.

As the external data and our DES Y3 observations are independent, we can simply use the posterior probability distributions from the external data as the prior probability for our analysis.

We have already used the PLANCK external data to justify the prior probability distributions for the so-called nuisance parameters in the Gower Street simulations (e.g. for  $\Omega_b$ ). Avoiding double counting of information, we simply use a new prior distribution only on the parameters of interest (i.e.  $S_8$ ,  $\Omega_m$ , and  $w$ ), which is given by the posterior distribution from the external analyses. As our original priors on the parameters of interest are flat, this requires a simple reweighting of the existing posterior distribution.

We do this in two ways: (i) relearning the posterior density of the parameters of interest from both our new DES Y3 MCMC chains and the MCMC chains from the external probes – we then multiply the learned densities; (ii) learning the posterior density of the parameters of interest for the external probes, and rerunning the MCMC chains using our learned DES Y3 simulation-based likelihood using the learned density as a prior. In both cases we use normalizing flows implemented in the package `tensimeter` to learn the densities from the existing posterior samples.

Both these methods give consistent results; we use method (i) in all quoted results.

	$\Lambda$ CDM DES+SNe+BAO	Planck 2018	Joint
$S_8$	$0.800 \pm 0.017$	$0.831 \pm 0.016$	$0.810 \pm 0.011$
$\Omega_m$	$0.290^{+0.010}_{-0.012}$	$0.3206 \pm 0.0096$	$0.3070 \pm 0.0067$
	$w$ CDM DES+SNe+BAO	Planck 2018	Joint
$S_8$	$0.807 \pm 0.020$	$0.840 \pm 0.016$	$0.813 \pm 0.010$
$\Omega_m$	$0.285 \pm 0.013$	$0.346^{+0.017}_{-0.024}$	$0.3106 \pm 0.0085$
$w$	$< -0.905$	$< -0.891$	$< -0.97$

Table IX. Cosmological constraints on parameters of interest in the  $\Lambda$ CDM and  $w$ CDM models. For each parameter we report the mean and 68 percent credible interval.

As we can see from Table IX the constraining power from late-time universe probes is comparable with that of CMB measurements. In  $\Lambda$ CDM this roughly amounts to a 2% constraint on  $S_8$  and a 3% constraint on  $\Omega_m$ . Since the two probes are independent the constraints add in quadrature almost perfectly, resulting in a 1% level joint constraint. The joint constraints are also shown in Fig. 6.

## VIII. SUMMARY

We presented a cosmological analysis using a combination of Gaussian and non-Gaussian statistics applied to the weak lensing mass (convergence) maps from the first three years (Y3) of the Dark Energy Survey (DES). Specifically, we considered: 1) second and third moments; 2) wavelet phase harmonics (WPH); 3) the scattering transform (ST). The second moments are Gaussian statistics, while the third moments probe additional non-Gaussian information of the fields. The WPH moments are the second moments of smoothed weak lensing mass maps that have undergone a non-linear transformation, enabling the exploration of the non-Gaussian features of the field. The ST coefficients are generated through a series of smoothing and modulus operations applied to the input field, followed by averaging. Both the WPH and ST are often linked to convolutional neural networks (CNNs) because their statistical definitions bear similarities to the architecture of CNNs. However, unlike CNNs, they do not require training data.

Our analysis is entirely based on simulations, spanning a space of seven  $w$ CDM cosmological parameters. It forward-models the most relevant sources of systematic errors in the data, including masks, noise variations, clustering of the sources, intrinsic alignments, and shear and redshift calibration. We have implemented a neural network to compress the summary statistics, and estimated the parameter posteriors using a simulation-based inference (SBI) approach. Our analysis setup has been extensively validated in paper I, which includes the validation of the mocks used in the analysis, empirical coverage tests of the posteriors obtained via simulation-based inference, tests on the impact of physical effects not included in the simulations (e.g. baryonic feedback effect),



and an end-to-end validation of the pipeline using an independent set of simulations. We complemented the tests from paper I with additional ones, particularly focusing on our data measurements. These include tests on the internal consistency of the different summary statistics measured in the data, assessments of goodness-of-fit, and B-mode null tests.

We demonstrated that incorporating and integrating various non-Gaussian statistics significantly enhances constraints compared to relying solely on Gaussian statistics (specifically, second moments). Notably, we achieved a 70 percent (90 percent) improvement in the Figure of Merit  $\text{FoM}(S_8, \Omega_m)$  for the  $\Lambda$ CDM ( $w$ CDM) model. The improvement in the  $\text{FoM}(S_8, \Omega_m)$  is a result of significant enhancements in both  $S_8$  and  $\Omega_m$ . By combining all summary statistics, we measured the amplitude of fluctuation parameter  $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5} = 0.794 \pm 0.017$  with 2 percent precision, under the assumption of a  $\Lambda$ CDM cosmology. In the case of a  $w$ CDM cosmology, the measurement was  $S_8 = 0.817 \pm 0.021$ . The inclusion of non-Gaussian statistics significantly tightened the constraints on  $A_{\text{IA}}$ , the amplitude of intrinsic alignment. Additionally, in the context of the  $w$ CDM scenario, these statistics also strengthened the constraints on the parameter  $w$ , obtaining  $w < -0.72$ , consistent with a  $\Lambda$ CDM scenario.

We compared our results with other weak lensing results from the DES Y3 data, finding good consistency. Our constraints outperform the results from the fiducial DES cosmic shear analysis (as expected) due to the extra information probed by the non-Gaussian statistics. We also find statistical agreement ( $< 2.2\sigma$ ) when comparing with results from external datasets: PLANCK constraints from the Cosmic Microwave Background, constraints from spectroscopic BAO measurements, and constraints from Type Ia supernovae.

### CONTRIBUTION STATEMENT

MG: Designed the project, led the analysis, and led the writing of the manuscript. GC: Performed comparisons with external datasets and edited the manuscript. NJ: Co-led the generation of the Gower Street simulations, proposed the simulation-based inference workflow, assisted with result interpretation, and edited the manuscript. LW: Co-led the generation of the Gower Street simulations, suggested mathematical notation, and edited the manuscript. AP: Recomputed DES Y3 cosmic shear posteriors using Gower Street priors. JP: Recomputed external dataset posteriors using Gower Street priors. JW: Helped perform measurements in simulations and helped with mocks testing. MR: Helped with interpretation of tension metrics, and with editing of the manuscript. BJ: Assisted with the interpretation of results and edited the manuscript. VA, GG, MY, CZ: Helped perform measurements in simulations. JB: Provided feedback on analyses choices and intrinsic alignment implementation. DA, SS, TK: DES internal

reviewers of the paper. MG, ES, AA, MB, MT, AC, CD, NM, ANA, IH, DG, GB, MJ, LFS, AF, JM, RPR, RC, CC, SP, IT, JP, JEP, CS: Production of the DES shape catalog. JM, AA, AA, CS, SE, JD, JM, DG, GB, MT, SD, AC, NM, BY, MR: Production of the DES redshift distribution. SE, BY, NK, EH, YZ: Production of the DES Balrog. MJ, GB, AA, CD, PFL, KB, IH, MG, AR: Production of the DES PSF. Builders: The remaining authors have made contributions to this paper that include, but are not limited to, the construction of DECam and other aspects of collecting the data; data processing and calibration; developing broadly used methods, codes, and simulations; running the pipelines and validation tests; and promoting the science analysis.

### ACKNOWLEDGEMENTS

The Gower Street simulations were generated under the DiRAC project p153 ‘Likelihood-free inference with the Dark Energy Survey’ (ACSP255/ACSC1) using DiRAC (STFC) HPC facilities ([www.dirac.ac.uk](http://www.dirac.ac.uk)).

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

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

OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include

ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciênciã e Tecnologia (INCT) do e-Universo (CNPq grant 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.

- 
- [1] Abbott, T. M. C., Aguena, M., Alarcon, A., et al., 2022, *Phys. Rev. D*, 105, 2, 023520, arXiv:2105.13549
  - [2] Ajani, V., Peel, A., Pettorino, V., Starck, J.-L., Li, Z., Liu, J., 2020, *Phys. Rev. D*, 102, 10, 103531
  - [3] Alam, S., Aubert, M., Avila, S., et al., 2021, *Phys. Rev. D*, 103, 8, 083533, arXiv:2007.08991
  - [4] Allys, E., Marchand, T., Cardoso, J. F., Villaescusa-Navarro, F., Ho, S., Mallat, S., 2020, *Phys. Rev. D*, 102, 10, 103506
  - [5] Alsing, J., Wandelt, B., Feeney, S., 2018, *MNRAS*, 477, 2874
  - [6] Amon, A., Gruen, D., Troxel, M. A., et al., 2022, *Phys. Rev. D*, 105, 2, 023514
  - [7] Anbajagane, D., Chang, C., Banerjee, A., et al., 2023, *MNRAS*, 526, 4, 5530, arXiv:2308.03863
  - [8] Aricò, G., Angulo, R. E., Hernández-Monteagudo, C., et al., 2020, *MNRAS*, 495, 4, 4800
  - [9] Asgari, M., Lin, C.-A., Joachimi, B., et al., 2021, *A&A*, 645, A104, arXiv:2007.15633
  - [10] Banerjee, A., Abel, T., 2023, *MNRAS*, 519, 4, 4856
  - [11] Barthelemy, A., Codis, S., Uhlemann, C., Bernardeau, F., Gavazzi, R., 2020, *MNRAS*, 492, 3, 3420
  - [12] Bishop, C. M., 1994, Mixture density networks, Tech. rep., Aston University
  - [13] Boruah, S. S., Lavaux, G., Hudson, M. J., 2022, *MNRAS*, 517, 3, 4529
  - [14] Boyle, A., Uhlemann, C., Friedrich, O., et al., 2021, *MNRAS*, 505, 2, 2886
  - [15] Bridle, S., King, L., 2007, *New Journal of Physics*, 9, 444
  - [16] Bruna, J., Mallat, S., 2013, arXiv e-prints, arXiv:1311.0407
  - [17] Chang, C., Pujol, A., Mawdsley, B., et al., 2018, *MNRAS*, 475, 3165
  - [18] Cheng, S., Marques, G. A., Grandón, D., et al., 2024, arXiv e-prints, arXiv:2404.16085, arXiv:2404.16085
  - [19] Cheng, S., Ting, Y.-S., Ménard, B., Bruna, J., 2020, *MNRAS*, 499, 4, 5902
  - [20] Cohen, A., Ryan, R. D., 1995, *Wavelets and multiscale signal processing*, Chapman & Hall
  - [21] Cordero, J. P., Harrison, I., Rollins, R. P., et al., 2022, *MNRAS*, 511, 2, 2170
  - [22] Dalal, R., Li, X., Nicola, A., et al., 2023, *Phys. Rev. D*, 108, 123519
  - [23] Dietrich, J. P., Hartlap, J., 2010, *MNRAS*, 402, 2, 1049
  - [24] Doux, C., Baxter, E., Lemos, P., et al., 2021, *MNRAS*, 503, 2, 2688
  - [25] Feldbrugge, J., van Engelen, M., van de Weygaert, R., Pranav, P., Vegter, G., 2019, *JCAP*, 2019, 9, 052
  - [26] Fluri, J., Kacprzak, T., Lucchi, A., et al., 2019, *Phys. Rev. D*, 100, 6, 063514
  - [27] Fluri, J., Kacprzak, T., Refregier, A., Amara, A., Lucchi, A., Hofmann, T., 2018, *Phys. Rev. D*, 98, 12, 123518
  - [28] Foreman-Mackey, D., Hogg, D. W., Lang, D., Goodman, J., 2013, *PASP*, 125, 306
  - [29] Gatti, M., Chang, C., Friedrich, O., et al., 2020, *MNRAS*, 498, 3, 4060
  - [30] Gatti, M., Giannini, G., Bernstein, G. M., et al., 2022, *MNRAS*, 510, 1, 1223
  - [31] Gatti, M., Jain, B., Chang, C., et al., 2022, *Phys. Rev. D*, 106, 8, 083509
  - [32] Gatti, M., Jeffrey, N., Whiteway, L., et al., 2024, *Phys. Rev. D*, 109, 063534
  - [33] Gatti, M., Jeffrey, N., Whiteway, L., et al., 2024, *MNRAS*, 527, 1, L115, arXiv:2307.13860
  - [34] Gatti, M., Sheldon, E., Amon, A., et al., 2021, *MNRAS*, 504, 3, 4312
  - [35] Górski, K. M., Hivon, E., Banday, A. J., et al., 2005, *ApJ*, 622, 2, 759, arXiv:astro-ph/0409513
  - [36] Harnois-Déraps, J., Martinet, N., Reischke, R., 2022, *MNRAS*, 509, 3, 3868
  - [37] Heavens, A. F., Jimenez, R., Lahav, O., 2000, *MNRAS*, 317, 965
  - [38] Heydenreich, S., Brück, B., Burger, P., et al., 2022, *A&A*, 667, A125
  - [39] Heydenreich, S., Brück, B., Harnois-Déraps, J., 2021, *A&A*, 648, A74
  - [40] Huff, E., Mandelbaum, R., 2017, arXiv e-prints, 1702.02600
  - [41] Jeffrey, N., Alsing, J., Lanusse, F., 2021, *MNRAS*, 501, 1, 954
  - [42] Jeffrey, N., Gatti, M., Chang, C., et al., 2021, *MNRAS*, 505, 3, 4626
  - [43] Jeffrey, N., Wandelt, B. D., 2020, *Third Workshop on Machine Learning and the Physical Sciences*, NeurIPS 2020, arXiv:2011.05991
  - [44] Jeffrey, N., Whiteway, L., Gatti, M., et al., 2024, arXiv e-prints, arXiv:2403.02314
  - [45] Kacprzak, T., Fluri, J., Schneider, A., Refregier, A., Stadel, J., 2023, *JCAP*, 2023, 2, 050
  - [46] Kacprzak, T., Kirk, D., Friedrich, O., et al., 2016, *MNRAS*, 463, 3653
  - [47] Kaiser, N., Squires, G., 1993, *ApJ*, 404, 441

- [48] Kratochvil, J. M., Haiman, Z., May, M., 2010, *Phys. Rev. D*, 81, 4, 043519
- [49] Kratochvil, J. M., Lim, E. A., Wang, S., Haiman, Z., May, M., Huffenberger, K., 2012, *Phys. Rev. D*, 85, 10, 103513
- [50] Li, S.-S., Hoekstra, H., Kuijken, K., et al., 2023, *A&A*, 679, A133, arXiv:2306.11124
- [51] Li, X., Zhang, T., Sugiyama, S., et al., 2023, *Phys. Rev. D*, 108, 123518
- [52] Liu, J., Petri, A., Haiman, Z., Hui, L., Kratochvil, J. M., May, M., 2015, *Phys. Rev. D*, 91, 6, 063507
- [53] Lu, T., Haiman, Z., Li, X., 2023, *MNRAS*, 521, 2, 2050
- [54] MacCrann, N., Becker, M. R., McCullough, J., et al., 2022, *MNRAS*, 509, 3, 3371, arXiv:2012.08567
- [55] Mallat, S., 1999, *A wavelet tour of signal processing*, Elsevier
- [56] Mallat, S., 2011, arXiv e-prints, arXiv:1101.2286
- [57] Mallat, S., et al., 2020, *Information and Inference: A Journal of the IMA*, 9, 3, 721, ISSN 2049-8772
- [58] Martinet, N., Schneider, P., Hildebrandt, H., et al., 2018, *MNRAS*, 474, 1, 712
- [59] Muir, J., Bernstein, G. M., Huterer, D., et al., 2020, *MNRAS*, 494, 3, 4454
- [60] Myles, J., Alarcon, A., Amon, A., et al., 2021, *MNRAS*, 505, 3, 4249
- [61] Papamakarios, G., Pavlakou, T., Murray, I., 2017, *Advances in neural information processing systems*, 30
- [62] Parroni, C., Cardone, V. F., Maoli, R., Scaramella, R., 2020, *A&A*, 633, A71
- [63] Parroni, C., Töllet, É., Cardone, V. F., Maoli, R., Scaramella, R., 2021, *A&A*, 645, A123
- [64] Peel, A., Pettorino, V., Giocoli, C., Starck, J.-L., Baldi, M., 2018, *A&A*, 619, A38
- [65] Petri, A., Liu, J., Haiman, Z., May, M., Hui, L., Kratochvil, J. M., 2015, *Phys. Rev. D*, 91, 10, 103511
- [66] Planck Collaboration, 2020, *A&A*, 641, A5, arXiv:1907.12875
- [67] Porqueres, N., Heavens, A., Mortlock, D., Lavaux, G., 2022, *MNRAS*, 509, 3, 3194
- [68] Porth, L., Smith, R. E., 2021, *MNRAS*, 508, 3, 3474
- [69] Potter, D., Stadel, J., Teyssier, R., 2017, *Computational Astrophysics and Cosmology*, 4, 1, 2
- [70] Raveri, M., Doux, C., 2021, *Phys. Rev. D*, 104, 4, 043504, arXiv:2105.03324
- [71] Raveri, M., Hu, W., 2019, *Phys. Rev. D*, 99, 4, 043506, arXiv:1806.04649
- [72] Ribli, D., Pataki, B. Á., Csabai, I., 2019, *Nature Astronomy*, 3, 93
- [73] Sánchez, C., Prat, J., Zacharegkas, G., et al., 2022, *Phys. Rev. D*, 105, 8, 083529, arXiv:2105.13542
- [74] Schneider, A., Teyssier, R., 2015, *JCAP*, 2015, 12, 049
- [75] Scolnic, D. M., Jones, D. O., Rest, A., et al., 2018, *ApJ*, 859, 2, 101, arXiv:1710.00845
- [76] Secco, L. F., Samuroff, S., Krause, E., et al., 2022, *Phys. Rev. D*, 105, 2, 023515
- [77] Shan, H., Liu, X., Hildebrandt, H., et al., 2018, *MNRAS*, 474, 1, 1116
- [78] Sheldon, E. S., Huff, E. M., 2017, *ApJ*, 841, 24
- [79] Thiele, L., Hill, J. C., Smith, K. M., 2020, *Phys. Rev. D*, 102, 12, 123545
- [80] Valogiannis, G., Dvorkin, C., 2022, *Phys. Rev. D*, 106, 10, 103509
- [81] Valogiannis, G., Dvorkin, C., 2022, *Phys. Rev. D*, 105, 10, 103534
- [82] Van Den Berg, J., 1999, *Wavelets in Physics*, Cambridge University Press
- [83] Van Waerbeke, L., Benjamin, J., Erben, T., et al., 2013, *MNRAS*, 433, 3373
- [84] Vicinanza, M., Cardone, V. F., Maoli, R., Scaramella, R., Er, X., 2016, arXiv e-prints, arXiv:1606.03892
- [85] Vicinanza, M., Cardone, V. F., Maoli, R., Scaramella, R., Er, X., 2018, *Phys. Rev. D*, 97, 2, 023519
- [86] Vicinanza, M., Cardone, V. F., Maoli, R., Scaramella, R., Er, X., Tereno, I., 2019, *Phys. Rev. D*, 99, 4, 043534
- [87] Zürcher, D., Fluri, J., Ajani, V., Fischbacher, S., Refregier, A., Kacprzak, T., 2023, *MNRAS*, 525, 1, 761, arXiv:2206.01450
- [88] Zürcher, D., Fluri, J., Sgier, R., Kacprzak, T., Refregier, A., 2021, *JCAP*, 2021, 1, 028
- [89] Zürcher, D., Fluri, J., Sgier, R., et al., 2022, *MNRAS*, 511, 2, 2075

## Appendix A: $\Lambda$ CDM validation

In this Appendix we perform a number of validation tests for the  $\Lambda$ CDM analysis, similar to what was done in Paper A for the  $w$ CDM analysis. In particular, we:

- conduct a baryonic contamination test, i.e. we verify that the scales used in the  $\Lambda$ CDM analysis are not affected by potential unmodelled baryonic feedback processes;
- test that additive biases due to PSF modelling errors were negligible at the data vector level;
- check that potential errors in our model for the source clustering effects were negligible;
- perform a coverage test to ensure that the size of the posteriors we recover is accurate;
- carry out an end-to-end  $\Lambda$ CDM analysis on an independent set of simulations (CosmogridV1 sims), demonstrating that we recover the true cosmology of the simulations.

Beyond these tests, we also add two sets of tests where:

- we show we can recover the true cosmology of the simulations even if the parameters describing redshift uncertainties and shear bias calibrations were  $2\sigma$  off their mean values;
- we show that our posteriors are not sensitive to the details of the simulations used (box size, number of particles, redshift resolution).

The initial test we conducted was the baryonic contamination test. This followed the same methodology as presented in Paper A, albeit tailored for a  $\Lambda$ CDM analysis. Our analysis does not include a model for the effects of baryons at small scales. Consequently, we need to ensure that the scales used are not compromised by baryonic effects. To this end, we employed a suite of simulations (CosmoGridV1) that have been post-processed to include baryonic feedback using the baryonic correction model [8, 74]. Each simulation is available in two

versions: one with baryonification and one without. We generate multiple instances of the DES maps using both versions to calculate our summary statistics. We then test whether the posteriors of cosmological parameters derived from data vectors including baryonic feedback do not exhibit significant bias when compared to those from data vectors without baryonic effects. We follow the criterion adopted by the DES Y3 cosmological analysis (Amon et al. 6; Secco & Samuroff et al. 76; Abbott et al. 1), which requires that the peak of the marginalized two-dimensional posterior of  $\Omega_m$  and  $S_8$  from the analysis of baryon-contaminated data must fall within  $0.3\sigma$  of that from clean data. Table X reports the contamination levels obtained using all available scales of our analyses, for all the individual summary statistics and certain combinations of them. Notably, all the summary statistics stayed within our pre-established limits for contamination, thereby showing the robustness of our analysis against potential baryonic feedback effects.

For the tests addressing potential PSF contamination, we emulated the methodology from paper I and analyzed two sets of mock data from the *CosmoGridV1* suite, i.e. with and without added contributions from PSF modelling errors as estimated in [34]. We confirmed that, even when focusing on a  $\Lambda$ CDM analysis, none of our summary statistic combinations showed a bias in the  $S_8$ - $\Omega_m$  plane larger than  $0.10\sigma$ . This indicates that PSF modelling errors are insignificant for the scale range considered in our study.

Regarding the tests for potential errors in modelling source clustering, we analyzed two sets of maps with assumed galaxy-matter biases of  $b = 0.5$  and  $b = 1.5$ , diverging from the standard value of unity used in our analysis. We found that for all our summary statistics combinations, the bias in the  $S_8$ - $\Omega_m$  plane was under  $0.10\sigma$ . Thus, the influence on cosmological parameters is minimal, confirming that our source clustering model is accurate enough for our analysis.

The next test aimed at ensuring that the confidence levels from simulation-based inference were correctly estimated. This involved carrying out an empirical coverage test. In such tests, the inference process is repeated multiple times to confirm that the estimated posteriors accurately reflect the true parameter probabilities. For our purposes, we ran the inference multiple (100) times, each time omitting one mock data vector (chosen at random) from the neural likelihood estimation. We specifically omitted mocks with  $w = -1$ , corresponding to the  $\Lambda$ CDM scenario, from the *DarkGridV1* simulation set. The likelihood for the excluded data vector was then calculated, the posterior determined, and its accuracy verified against the known parameter values. We assessed coverage probabilities within the three-dimensional parameter space of  $\Omega_m$ ,  $S_8$ ,  $A_{1A}$  using the TARP package, which applies the ‘Tests of Accuracy with Random Points’ (TARP) method for estimating the coverage probabilities of generative posterior estimators. The outcomes, illustrated in Fig. 7, reveal that the expected coverage corresponds with the credibility levels

at 5 percent, signifying that our posterior estimates are properly calibrated.

We then confirm our pipeline’s ability to accurately recover the actual cosmology from a set of simulations not previously used in its construction. For this purpose, we use 400 independent DES Y3 mock catalogues, generated by the *CosmoGridV1* simulations, identical to those employed for the  $w$ CDM analysis validation in paper I. None of these *CosmoGridV1* simulations are used during the training of our inference pipeline (i.e. during the compression or when training the NDEs), but they are rather used as a ‘target’ data vector. Each mock has the same cosmology; we further assume no intrinsic alignment, while for the other nuisance parameters (shear calibration and redshift uncertainties) we assume values at the centre of the priors. By measuring all summary statistics across the mocks and taking their average, we mitigate the influence of noise. The resulting posterior distributions for the cosmological parameters  $\Omega_m$  and  $S_8$ , derived from the second moments and the combination of all summary statistics, are depicted in Fig. 8. The simulated analysis recovers the true cosmological parameters of the simulations, underscoring the reliability of our analytical framework.

In the final series of tests, we evaluate our ability to accurately recover the cosmological parameters of simulations under various conditions. Specifically, we assess whether the correct cosmology can be identified when: 1) the simulation includes shear biases that deviate by  $2\sigma$  from the average values of their prior distributions; 2) the redshift mean values diverge by  $2\sigma$  from their prior averages; 3) the simulations use double the number of particles; 4) the simulations have been produced with double the box size; and 5) the simulations feature twice the redshift resolution. For these latter three scenarios, we used benchmark simulations from the *CosmoGridV1* suite. The results, depicted in Fig. 10, confirm that in every instance we successfully retrieved the true cosmological values of the simulations.

## Appendix B: B-modes test

At first order, weak lensing is not expected to produce B-modes. Consequently, B-modes serve as a null test for identifying potential systematic effects not accounted for in the analysis. However, due to our map-making procedure using the Kaiser-Squires algorithm, B-modes can emerge as a result of masking effects, as noted [31]. This occurs because a small fraction of the E-mode power leaks into the B-mode map, predominantly at large scales. In our data vector, we included the second moments and WHPG of the B-mode maps, although their contribution to the cosmological constraints is small. We opted not to include other summary statistics for B-modes, as this would double the length of the data vector, complicating the data compression without significantly improving our constraints. Nevertheless, we conducted a comparison of B-modes in our data

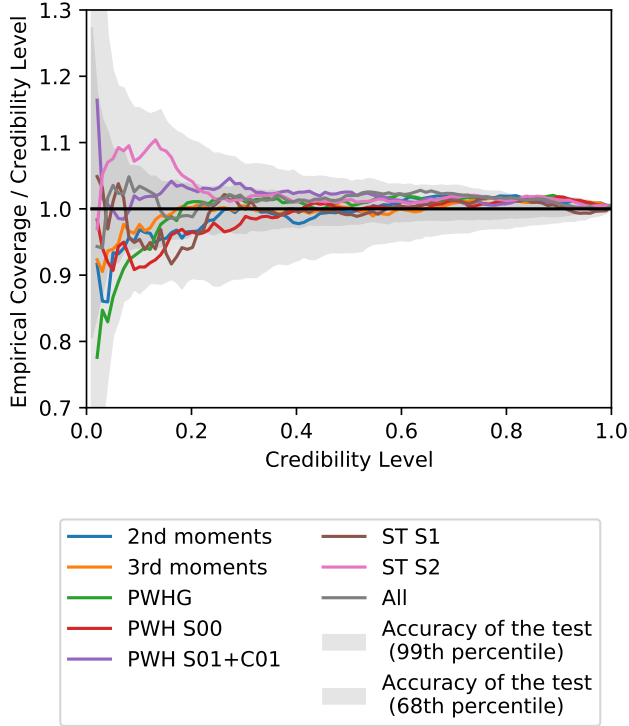


Figure 7. Expected coverage probability of posteriors derived from the simulation-based inference pipeline, using various summary statistics, with respect to the credibility level. The grey shaded areas represent the precision of the test, constrained by the finite number of posteriors used in this analysis.

Table X. Bias in the parameter posteriors assessed by contrasting the results from an analysis that simulated mocks with baryonic feedback against one using mocks with no such feedback. The impact on various summary statistics are quantified by measuring the separation between the peaks of the posterior distributions in the  $S_8$ - $\Omega_m$  plane. We found that all biases are well below the  $0.3\sigma$  threshold (the upper limit of bias tolerated in our analysis).

Summary Statistic(s)	Contamination $S_8 - \Omega_m$
2nd moments	$0.06\sigma$
WPHG	$0.05\sigma$
3rd moments	$0.05\sigma$
WPH S00	$0.01\sigma$
WPH S01+C01	$0.05\sigma$
WPH S00+S01+C01	$0.11\sigma$
ST1	$0.05\sigma$
ST2	$0.03\sigma$
ST1+ST2	$0.04\sigma$
2nd+3rd moments	$0.10\sigma$
2nd moments+WPH	$0.05\sigma$
2nd moments+ST1	$0.08\sigma$
2nd+3rd moments+ST+WPH	$0.15\sigma$

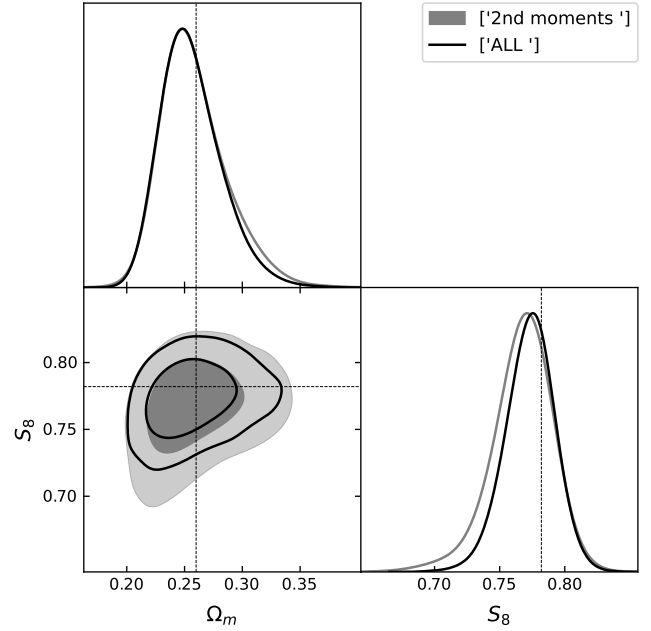


Figure 8. Posterior distributions of the cosmological parameters  $\Omega_m$  and  $S_8$  for different second moments and for the combinations of all the summary statistics considered in this work, as measured in *CosmoGridV1* simulations. The dotted black lines indicate the values of the cosmological parameters in the simulations. The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions.

against those measured in simulations for the summary statistics beyond second moments and WPHG. Such a comparison is illustrated in Figure 9. Specifically, we compared against measurements from the *CosmoGridV1* simulations, which were performed at a fixed cosmology. The measurements from our data and the simulations show good agreement. We did not make a significant detection of B-modes for third moments and WPH S01. These results on third moments corroborate the findings of [31]. For other statistics (ST1, ST2, and WPH S00), both data and simulations indicate a slight, non-null B-mode signal at large scales, likely attributable to E-mode leakage caused by masking effects. Given the strong correspondence between simulations and data, we conclude that our B-mode analysis reveals no evidence of unaccounted systematics in our simulations.

### Appendix C: NDEs and parameters posterior

In this work we used four different neural density estimators (NDEs) to estimate the posteriors. In particular, we used two different Gaussian Mixture Density Networks (MDNs) and two different Masked Autoencoders for Distribution Estimation (MADEs). Whenever we showed a posterior or reported the constraints on some parameters in this work, we always obtained these by stacking the four different NDEs. Assuming all

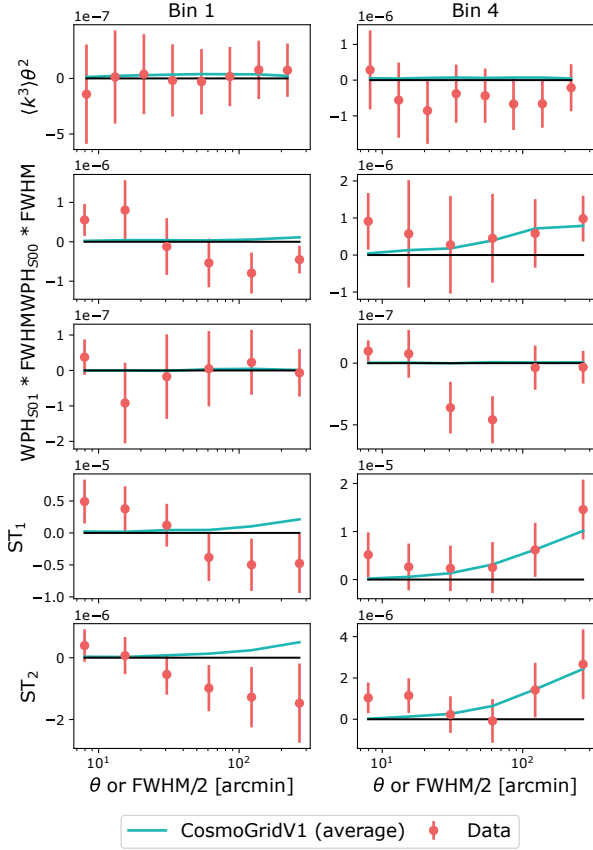


Figure 9. Comparison of the measured summary statistics of B-mode convergence maps in our data with those obtained from the `CosmoGridV1` simulations. We did not include measurements for second moments and WPHG in this comparison, as these statistics are already incorporated into the data vectors used for the cosmological analysis.

the NDEs are flexible enough to describe our likelihood surface, they should all agree in the limit in which the number of simulations used for training becomes large. Fig. 11 shows the posteriors obtained by each individual NDE for our most constraining case (i.e. the combination of all summary statistics); the posteriors are very similar, indicating that our posteriors estimates are robust and stable.

#### Appendix D: Impact of the `DarkGridV1` simulation suite

Our analysis uses simulations in two distinct ways: for training the compression algorithm and for learning the likelihood surface with the NDEs. In this work, we have expanded the simulation suite used in our validation study (paper I), by including simulations from the `DarkGridV1` suite. These additional simulations are specific to the  $\Lambda$ CDM model, meaning they do not vary the pa-

rameter  $w$ . Furthermore, they do not vary other parameters such as  $h_{100}$ ,  $n_s$ ,  $m$ ,  $\Omega_b$ , or the neutrino mass (unlike those in the Gower Street suite). Our NDEs do not

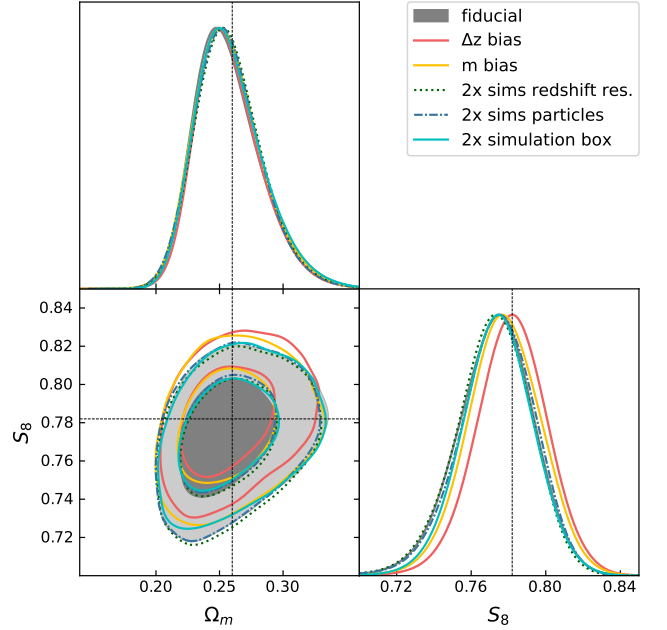


Figure 10. Posterior distributions of the cosmological parameters  $\Omega_m$  and  $S_8$  for different combinations of all the summary statistics considered in this work, as measured in a number of different `CosmoGridV1` simulations.

explicitly learn the dependence on these latter parameters; instead, these parameters are effectively marginalized over according to the distribution followed by the simulations. Incorporating the `DarkGridV1` suite simulations thus modifies the distribution of mocks, which could impact the posterior. Figure 12 shows a comparison of our posterior on the data for the  $\Lambda$ CDM case, with and without including the `DarkGridV1` suite in learning the likelihood surface. The posteriors are very similar, indicating that the impact of the differing parameter distribution is negligible.

#### Appendix E: Cosmological constraints of individual statistics

We show in this Appendix the posteriors for  $\Omega_m$  and  $S_8$  obtained from individual statistics, i.e. second moments, third moments, WPHG, WPH S01, WPH S00, ST1, and ST2, and their combination. The posteriors are shown in Fig. 13, for both  $w$ CDM (left) and  $\Lambda$ CDM (right). Constraints from individual probes are largely overlapping and consistent with each other. Third moments and WPH S01 are the less constraining statistic among all those explored here; this is partially because all the other non-Gaussian statistics (ST1, ST2 and WPH S00) also probe part of the Gaussian information of the field, and hence are more constraining.

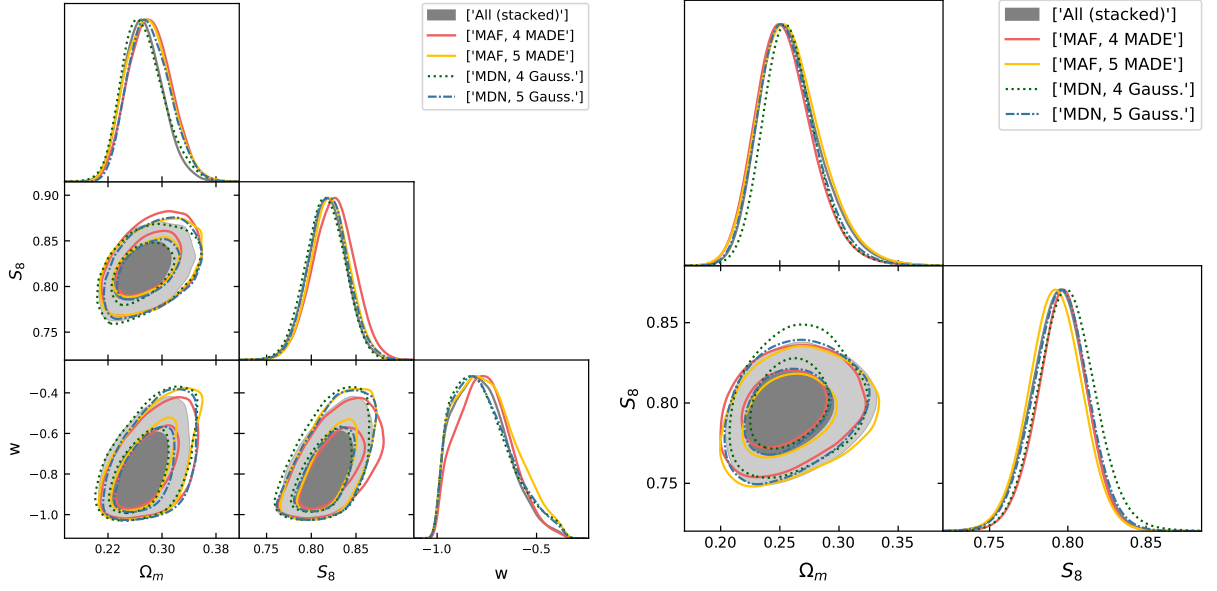


Figure 11. Posterior distributions of the cosmological parameters  $\Omega_m$  and  $\sigma_8$  for the combination of all the summary statistics considered in this work, as measured in data, for the  $\Lambda$ CDM (right) and  $w$ CDM (left) analyses. We show the different posteriors as estimated by the different NDEs used in this work; we also show their stacked combination (the fiducial setup used in the other Figures of this paper). The dotted black lines indicate the values of the cosmological parameters in the simulations. The two-dimensional marginalised contours in these figures show the 68 percent and 95 percent credible regions.

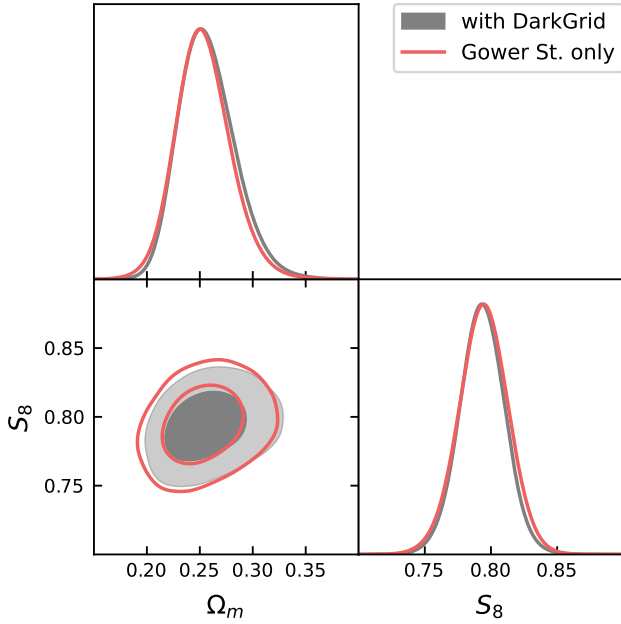


Figure 12. Posterior distributions of the cosmological parameters  $\Omega_m$  and  $\sigma_8$  obtained including or not `DarkGridV1` simulations into our analysis setup. The posteriors are shown for the  $\Lambda$ CDM scenario, and for the case where all the summary statistics are combined together.

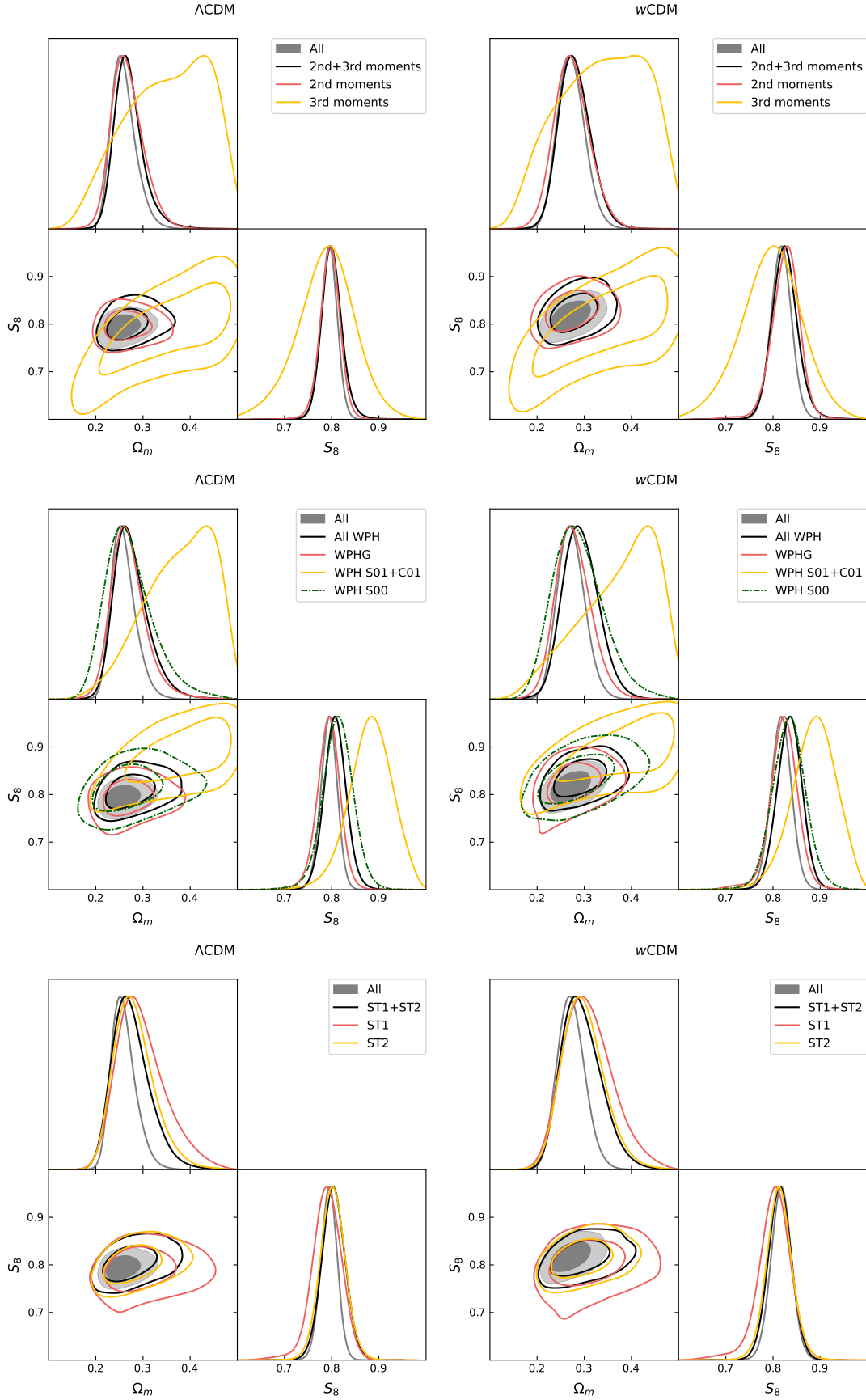


Figure 13. Posterior distributions of the cosmological parameters  $\Omega_m$  and  $S_8$  from individual statistics and their combination, for both  $w$ CDM (right) and  $\Lambda$ CDM (left).