

## Dark Energy Survey Year 3 Results: Cosmology from Cosmic Shear and Robustness to Modeling Uncertainty

L. F. Secco,<sup>1,2,\*</sup> S. Samuroff,<sup>3,4,†</sup> E. Krause,<sup>5</sup> B. Jain,<sup>2</sup> J. Blazek,<sup>4,6</sup> M. Raveri,<sup>2</sup> A. Campos,<sup>3</sup> A. Amon,<sup>7</sup> A. Chen,<sup>8</sup> C. Doux,<sup>2</sup> A. Choi,<sup>9,10</sup> D. Gruen,<sup>7,11,12</sup> G. M. Bernstein,<sup>2</sup> C. Chang,<sup>1</sup> J. DeRose,<sup>13</sup> J. Myles,<sup>7,11,12</sup> A. Ferté,<sup>14</sup> P. Lemos,<sup>15</sup> D. Huterer,<sup>8</sup> J. Prat,<sup>1</sup> M. A. Troxel,<sup>16</sup> N. MacCrann,<sup>17</sup> A. R. Liddle,<sup>18,19,20</sup> T. Kacprzak,<sup>21</sup> X. Fang,<sup>5</sup> C. Sánchez,<sup>2</sup> S. Pandey,<sup>2</sup> S. Dodelson,<sup>3</sup> P. Chintalapati,<sup>22</sup> K. Hoffmann,<sup>23</sup> A. Alarcon,<sup>24</sup> O. Alves,<sup>25</sup> F. Andrade-Oliveira,<sup>25,26</sup> E. J. Baxter,<sup>27</sup> K. Bechtol,<sup>28</sup> M. R. Becker,<sup>24</sup> A. Brandao-Souza,<sup>29,30</sup> H. Camacho,<sup>25,26</sup> A. Carnero Rosell,<sup>31,32,33</sup> M. Carrasco Kind,<sup>34,35</sup> R. Cawthon,<sup>28</sup> J. P. Cordero,<sup>36</sup> M. Crocce,<sup>37,38</sup> C. Davis,<sup>7</sup> E. Di Valentino,<sup>36</sup> A. Drlica-Wagner,<sup>39,40,1</sup> K. Eckert,<sup>2</sup> T. F. Eifler,<sup>5,14</sup> M. Elidaiana,<sup>8</sup> F. Elsner,<sup>41</sup> J. Elvin-Poole,<sup>9,10</sup> S. Everett,<sup>42</sup> P. Fosalba,<sup>37,38</sup> O. Friedrich,<sup>43</sup> M. Gatti,<sup>2</sup> G. Giannini,<sup>44</sup> R. A. Gruendl,<sup>34,35</sup> I. Harrison,<sup>45,36</sup> W. G. Hartley,<sup>46</sup> K. Herner,<sup>40</sup> H. Huang,<sup>5</sup> E. M. Huff,<sup>14</sup> M. Jarvis,<sup>2</sup> N. Jeffrey,<sup>47,48</sup> N. Kuropatkin,<sup>40</sup> P. -F. Leget,<sup>49,50,51</sup> J. Muir,<sup>7</sup> J. McCullough,<sup>7,11,12</sup> A. Navarro Alsina,<sup>52,26</sup> Y. Omori,<sup>39,1,7</sup> Y. Park,<sup>53</sup> A. Porredon,<sup>9,10</sup> R. Rollins,<sup>36</sup> A. Roodman,<sup>7,12</sup> R. Rosenfeld,<sup>54,26</sup> A. J. Ross,<sup>55</sup> E. S. Rykoff,<sup>7,12</sup> J. Sanchez,<sup>40</sup> I. Sevilla-Noarbe,<sup>31</sup> E. S. Sheldon,<sup>56</sup> T. Shin,<sup>2</sup> I. Tutusaus,<sup>38</sup> T. N. Varga,<sup>41,57</sup> N. Weaverdyck,<sup>8</sup> R. H. Wechsler,<sup>11,7,12</sup> B. Yanny,<sup>58</sup> B. Yin,<sup>3</sup> Y. Zhang,<sup>40</sup> J. Zuntz,<sup>59</sup> T. M. C. Abbott,<sup>60</sup> M. Aguena,<sup>30</sup> S. Allam,<sup>40</sup> J. Annis,<sup>40</sup> D. Bacon,<sup>61</sup> E. Bertin,<sup>62,63</sup> S. Bhargava,<sup>64</sup> S. L. Bridle,<sup>36</sup> D. Brooks,<sup>47</sup> E. Buckley-Geer,<sup>39,40</sup> D. L. Burke,<sup>7,12</sup> J. Carretero,<sup>65</sup> M. Costanzi,<sup>66,67,68</sup> L. N. da Costa,<sup>30,69</sup> J. De Vicente,<sup>70</sup> H. T. Diehl,<sup>40</sup> J. P. Dietrich,<sup>71</sup> P. Doel,<sup>47</sup> I. Ferrero,<sup>72</sup> B. Flaugher,<sup>40</sup> J. Frieman,<sup>40,1</sup> J. García-Bellido,<sup>73</sup> E. Gaztanaga,<sup>37,38</sup> D. W. Gerdes,<sup>74,8</sup> T. Giannantonio,<sup>43,75</sup> J. Gschwend,<sup>30,69</sup> G. Gutierrez,<sup>40</sup> S. R. Hinton,<sup>76</sup> D. L. Hollowood,<sup>42</sup> K. Honscheid,<sup>9,10</sup> B. Hoyle,<sup>71,77</sup> D. J. James,<sup>78</sup> T. Jeltema,<sup>42</sup> K. Kuehn,<sup>79,80</sup> O. Lahav,<sup>47</sup> M. Lima,<sup>81,30</sup> H. Lin,<sup>40</sup> M. A. G. Maia,<sup>30,69</sup> J. L. Marshall,<sup>82</sup> P. Martini,<sup>9,83,84</sup> P. Melchior,<sup>85</sup> F. Menanteau,<sup>86,34</sup> R. Miquel,<sup>87,65</sup> J. J. Mohr,<sup>71,77</sup> R. Morgan,<sup>28</sup> R. L. C. Ogando,<sup>30,69</sup> A. Palmese,<sup>40,1</sup> F. Paz-Chinchón,<sup>86,43</sup> D. Petravick,<sup>86</sup> A. Pieres,<sup>30,69</sup> A. A. Plazas Malagón,<sup>85</sup> M. Rodriguez-Monroy,<sup>70</sup> A. K. Romer,<sup>64</sup> E. Sanchez,<sup>70</sup> V. Scarpine,<sup>40</sup> M. Schubnell,<sup>8</sup> D. Scolnic,<sup>16</sup> S. Serrano,<sup>37,38</sup> M. Smith,<sup>88</sup> M. Soares-Santos,<sup>8</sup> E. Suchyta,<sup>89</sup> M. E. C. Swanson,<sup>86</sup> G. Tarle,<sup>8</sup> D. Thomas,<sup>61</sup> and C. To<sup>11,7,12</sup>

(DES Collaboration)

<sup>1</sup>Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

<sup>2</sup>Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>3</sup>McWilliams Center for Cosmology, Department of Physics,  
Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>4</sup>Department of Physics, Northeastern University, Boston, MA, 02115, USA

<sup>5</sup>Department of Astronomy/Steward Observatory, University of Arizona,  
933 North Cherry Avenue, Tucson, AZ 85721-0065, USA

<sup>6</sup>Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), 1290 Versoix, Switzerland

<sup>7</sup>Kavli Institute for Particle Astrophysics & Cosmology,  
P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

<sup>8</sup>Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

<sup>9</sup>Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA

<sup>10</sup>Department of Physics, The Ohio State University, Columbus, OH 43210, USA

<sup>11</sup>Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

<sup>12</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

<sup>13</sup>Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA

<sup>14</sup>Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

<sup>15</sup>Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

<sup>16</sup>Department of Physics, Duke University Durham, NC 27708, USA

<sup>17</sup>Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK

<sup>18</sup>Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, 1769-016 Lisboa, Portugal

<sup>19</sup>Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK

<sup>20</sup>Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada

<sup>21</sup>Institute for Particle Physics and Astrophysics, ETH Zürich,  
Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland

<sup>22</sup>Department of Physics, Northern Illinois University, DeKalb, IL 60115, USA

<sup>23</sup>Institute for Computational Science, University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland

<sup>24</sup>Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

<sup>25</sup>Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil

<sup>26</sup>Laboratório Interinstitucional de e-Astronomia - LIneA,  
Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

<sup>27</sup>Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

<sup>28</sup>Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, WI 53706-1390

<sup>29</sup>Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil

- <sup>30</sup>Laboratório Interinstitucional de e-Astronomia - LIneA,  
Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- <sup>31</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), E-28040 Madrid, Spain
- <sup>32</sup>Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
- <sup>33</sup>Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
- <sup>34</sup>Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
- <sup>35</sup>National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA
- <sup>36</sup>Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy,  
University of Manchester, Oxford Road, Manchester, M13 9PL, UK
- <sup>37</sup>Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
- <sup>38</sup>Institute of Space Sciences (ICE, CSIC), Campus UAB,  
Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
- <sup>39</sup>Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA
- <sup>40</sup>Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
- <sup>41</sup>Max Planck Institute für Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany
- <sup>42</sup>Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
- <sup>43</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- <sup>44</sup>Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of  
Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain
- <sup>45</sup>Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
- <sup>46</sup>Département de Physique Théorique and Center for Astroparticle Physics,  
Université de Genève, 24 quai Ernest Ansermet, CH-1211 Geneva, Switzerland
- <sup>47</sup>Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
- <sup>48</sup>Laboratoire de Physique de l'Ecole Normale Supérieure, ENS,  
Université PSL, CNRS, Sorbonne Université, Université de Paris, Paris, France
- <sup>49</sup>Université Clermont Auvergne, CNRS/IN2P3, Laboratoire de Physique de Clermont, F-63000 Clermont-Ferrand, France
- <sup>50</sup>Kavli Institute for Particle Astrophysics & Cosmology,  
Department of Physics, Stanford University, Stanford, CA 94305
- <sup>51</sup>LPNHE, CNRS/IN2P3, Sorbonne Université, Paris Diderot,  
Laboratoire de Physique Nucléaire et de Hautes Énergies, F-75005, Paris, France
- <sup>52</sup>Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil
- <sup>53</sup>Kavli Institute for the Physics and Mathematics of the Universe (WPI),  
UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
- <sup>54</sup>ICTP South American Institute for Fundamental Research Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil
- <sup>55</sup>Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, Ohio, USA
- <sup>56</sup>Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
- <sup>57</sup>Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany
- <sup>58</sup>Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
- <sup>59</sup>Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK
- <sup>60</sup>Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile
- <sup>61</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
- <sup>62</sup>CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France
- <sup>63</sup>Sorbonne Universités, UPMC Univ Paris 06, UMR 7095,  
Institut d'Astrophysique de Paris, F-75014, Paris, France
- <sup>64</sup>Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, UK
- <sup>65</sup>Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of  
Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain
- <sup>66</sup>Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy
- <sup>67</sup>INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy
- <sup>68</sup>Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy
- <sup>69</sup>Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- <sup>70</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
- <sup>71</sup>Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany
- <sup>72</sup>Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway
- <sup>73</sup>Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, 28049 Madrid, Spain
- <sup>74</sup>Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
- <sup>75</sup>Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- <sup>76</sup>School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia
- <sup>77</sup>Max Planck Institute für Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany
- <sup>78</sup>Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
- <sup>79</sup>Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia
- <sup>80</sup>Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA
- <sup>81</sup>Departamento de Física Matemática, Instituto de Física,  
Universidade de São Paulo, CP 66318, São Paulo, SP, 05314-970, Brazil

<sup>82</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>83</sup>*Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*

<sup>84</sup>*Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA 02138*

<sup>85</sup>*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*

<sup>86</sup>*Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*

<sup>87</sup>*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

<sup>88</sup>*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*

<sup>89</sup>*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831*

(Dated: January 17, 2022)

This work and its companion paper, [Amon et al. \(2021\)](#), present cosmic shear measurements and cosmological constraints from over 100 million source galaxies in the Dark Energy Survey (DES) Year 3 data. We constrain the lensing amplitude parameter  $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$  at the 3% level in  $\Lambda$ CDM:  $S_8 = 0.759^{+0.025}_{-0.023}$  (68% CL). Our constraint is at the 2% level when using angular scale cuts that are optimized for the  $\Lambda$ CDM analysis:  $S_8 = 0.772^{+0.018}_{-0.017}$  (68% CL). With cosmic shear alone, we find no statistically significant constraint on the dark energy equation-of-state parameter at our present statistical power. We carry out our analysis blind, and compare our measurement with constraints from two other contemporary weak lensing experiments: the Kilo-Degree Survey (KiDS) and Hyper-Suprime Camera Subaru Strategic Program (HSC). We additionally quantify the agreement between our data and external constraints from the Cosmic Microwave Background (CMB). Our DES Y3 result under the assumption of  $\Lambda$ CDM is found to be in statistical agreement with Planck 2018, although favors a lower  $S_8$  than the CMB-inferred value by  $2.3\sigma$  (a  $p$ -value of 0.02). This paper explores the robustness of these cosmic shear results to modeling of intrinsic alignments, the matter power spectrum and baryonic physics. We additionally explore the statistical preference of our data for intrinsic alignment models of different complexity. The fiducial cosmic shear model is tested using synthetic data, and we report no biases greater than  $0.3\sigma$  in the plane of  $S_8 \times \Omega_m$  caused by uncertainties in the theoretical models.

## I. INTRODUCTION

Discoveries and advances in modern cosmology have resulted in a remarkably simple standard cosmological model, known as  $\Lambda$ CDM. The model is specified by a spatially flat universe, governed by the general theory of relativity, which contains baryonic matter, dark matter, and a dark energy component that causes the expansion of the Universe to accelerate. Although remarkably simple, it appears to be sufficient to describe a great many observations, including the stability of cold disk galaxies, flat galaxy rotation curves, observations of strong gravitational lensing in clusters, the acceleration of the expansion of the Universe as inferred by type Ia supernovae (SNe Ia), and the pattern of temperature fluctuations in the Cosmic Microwave Background (CMB). Yet despite all this,  $\Lambda$ CDM is fundamentally mysterious in the sense that the physical nature of its two main components, dark matter and dark energy, is still completely unknown.

The success of  $\Lambda$ CDM has, however, been shaken in recent years by new experimental results. We have seen tentative hints that the model might fail to simultaneously describe the late- (low redshift) and early-time (high redshift) Universe. To take one prominent example, constraints on the local expansion parameter  $H_0$  obtained from the local distance ladder and SNe Ia appear to be in tension with those inferred by the CMB ([Planck Collaboration 2020b](#)) at a statistically significant level ([Riess et al. 2021](#)), with varying levels of significance being reported by different probes ([Alam et al. 2021](#), [Freedman et al.](#)

[2019](#)). In a separate but analogous tension, the value of the  $S_8 \equiv \sigma_8 (\Omega_m/0.3)^{1/2}$  parameter — the amplitude of mass fluctuations  $\sigma_8$  scaled by the square root of matter density  $\Omega_m$  — differs when inferred via cosmological lensing ([Asgari et al. 2021](#), [Hikage et al. 2019](#), [Troxel et al. 2018](#)) from the value obtained using Planck (assuming  $\Lambda$ CDM; [Planck Collaboration 2020b](#)) at the level of  $2 - 3\sigma$ . Other probes of the late Universe, in particular spectroscopic galaxy clustering ([Tröster et al. 2020](#)), redshift-space distortions ([Alam et al. 2021](#)) and the abundance of galaxy clusters ([Dark Energy Survey Collaboration 2020](#), [Mantz et al. 2015](#)), also all tend to prefer relatively low values of  $S_8$ . Although the evidence is by no means definitive, we are perhaps beginning to see hints of new physics, and so stress-testing  $\Lambda$ CDM with new measurements is extremely important.

Cosmic shear, or cosmological weak lensing (the two-point correlation function of gravitational shear), is one of the most informative of the low redshift probes. It has two main advantages, as a means to infer the properties of the large scale Universe ([Bartelmann & Schneider 2001](#), [Frieman et al. 2008](#), [Hu & Jain 2004](#), [Huterer 2002](#)). First, the signal is insensitive to galaxy bias, which is a significant source of uncertainty in cosmological analyses based on galaxy clustering and galaxy-galaxy lensing. Second, weak lensing is sensitive both to the *geometry* of the Universe through the lensing kernel (which is a function of  $H_0$  and ratios of angular diameter distances), and also to the *growth of structure* and its evolution in redshift. Since geometry and structure growth are tightly related to the evolution of dark energy and its equation-of-state parameter  $w$ , this sensitivity carries over to the cosmic shear signal.

Cosmic shear was first measured over twenty years ago, roughly simultaneously by a number of groups ([Bacon et al.](#)

\* secco@uchicago.edu

† s.samuroff@northeastern.edu

2000, Kaiser et al. 2000, Van Waerbeke et al. 2000, Wittman et al. 2000). Although too noisy to constrain cosmological parameters, these observations represented the first steps towards fulfilling the potential pointed out by theoretical studies years earlier (Hu 1999, Jain & Seljak 1997, Miralda-Escude 1991). The intervening two decades have seen steady improvements in signal-to-noise and cosmological constraining power, as new ground- and space-based lensing data sets have become available (Asgari et al. 2017, Asgari et al. 2021, Benjamin et al. 2007, Brown et al. 2003, Erben et al. 2013, Fu et al. 2008, Hamana et al. 2003, Hamana et al. 2020, Hetterscheidt et al. 2007, Heymans et al. 2005, Hikage et al. 2019, Hoekstra et al. 2002, Huff et al. 2014a,b, Jarvis et al. 2003, 2006, Jee et al. 2013, 2016, Kilbinger et al. 2013, Kitching et al. 2014, Leauthaud et al. 2007, Lin et al. 2012, Massey et al. 2007a, Miller et al. 2013, Refregier et al. 2002, Rhodes et al. 2004, Schrabback et al. 2010, Yoon et al. 2019). As the volume and quality of lensing data have improved, so too have the methods used to study it, with the development of an array of sophisticated statistical and theoretical tools. There has, for example, been a coherent effort to test and improve shape measurement algorithms using increasingly complex image simulations (Bridle et al. 2010, Heymans et al. 2006, Kitching et al. 2012, Mandelbaum et al. 2014, Massey et al. 2007b). Methods for estimating the distribution of source galaxies along the line of sight have also gradually evolved to become highly sophisticated, incorporating various sources of information (Gatti & Vielzeuf et al., 2018; Prat & Baxter et al., 2019; Buchs, Davis et al. 2019; Alarcon et al. 2020, Sánchez & Bernstein 2019, Wright et al. 2020).

Alongside the Dark Energy Survey (DES)<sup>1</sup>, the major lensing surveys of the current generation are the Kilo-Degree Survey (KiDS; de Jong et al. 2013)<sup>2</sup> and the Hyper-Suprime Camera Subaru Strategic Program (HSC; Aihara et al. 2018)<sup>3</sup>. We show the approximate, nominal footprints of these surveys in Fig. 1. Each of these three collaborations have, in recent years, released cosmic shear analyses analogous to the one presented in this paper. Lensing analyses based on HSC data were carried out over a footprint of 136.9 deg<sup>2</sup> split into six fields (red patches in Fig. 1; Mandelbaum et al. 2018a); they presented consistent cosmology results using two types of statistics: real space correlation functions (Hamana et al. 2020) and harmonic space power spectra (Hikage et al. 2019). More recently, the KiDS collaboration released results based on approximately 1000 deg<sup>2</sup> of data (blue patches in Fig. 1; Giblin et al. 2021), and presented an analysis of band-power spectra, correlation functions and the complete orthogonal sets of E-/B-mode integrals (COSEBIs Asgari et al. 2021). They further combined their cosmic shear results with external spectroscopic data from BOSS (Alam et al. 2015) to obtain a 3×2pt constraint (Heymans et al. 2021), which is internally consistent with their cosmic shear results, but differs from Planck in the full parameter space by  $\sim 2\sigma$ .

The trend observed in earlier cosmic shear studies is that the amplitude of the cosmic shear signal (tied to the amplitude of matter fluctuations through the  $S_8$  parameter) is lower than that extrapolated from the CMB. In order to demonstrate whether this discrepancy is physical and significant, we must have a high degree of confidence in our modeling of the data and its possible systematic errors. Among the most significant of these sources of systematic error are intrinsic alignments (IAs), or astrophysically sourced correlations of galaxy shapes, which mimic cosmic shear. Given how difficult it is to disentangle IAs from lensing, the most common approach is to forward-model their effect, assuming a model for the IA power spectrum with a number of free parameters. Depending on the galaxy sample, however, IA model insufficiency can easily translate into a bias in cosmological parameters (Blazek et al. 2019, Krause et al. 2016). In addition to IAs, effects such as nonlinear growth and the impact of baryons on the large-scale distribution of dark matter can alter the matter power spectrum in a significant way, and so bias the inferred lensing amplitude if neglected (DeRose et al. 2019b, Huang et al. 2021, Martinelli et al. 2021, Schneider et al. 2019, Yoon & Jee 2021). Although it is clear that these effects are scale-dependent, finding the angular scales where our modeling is sufficient is by no means straightforward. This paper describes the choices made in modeling and scale cuts, and validates that the potential biases on cosmological parameters are smaller than the statistical uncertainties.

Our companion paper (Amon et al. 2021) presents a detailed investigation of observational errors that can similarly bias cosmological inference. Undiagnosed biases in the shear measurement process, for example, can lead one to incorrectly infer the lensing amplitude. Likewise, errors in the estimation of galaxy redshift distributions  $n(z)$  can subtly alter the interpretation of the lensing measurement, both in terms of cosmology and of IAs. Amon et al. (2021) demonstrate that these measurement systematic errors are well controlled in the Y3 cosmic shear analysis. *We note that the main cosmological constraints presented in both papers are identical.*

The cosmic shear analysis presented in this paper, and in Amon et al. (2021), is part of a series of Year 3 cosmological results from large-scale structure produced by the Dark Energy Survey Collaboration. This work relies on many companion papers that validate the data, catalogs and theoretical methods; those papers, as well as this one, feed into the main “3 × 2pt” constraints, which combine cosmic shear with galaxy-galaxy lensing and galaxy clustering in  $\Lambda$ CDM and  $w$ CDM (Dark Energy Survey Collaboration 2021a), as well as extended cosmological parameter spaces (Dark Energy Survey Collaboration 2021b). These include:

- The construction and validation of the GOLD catalog of objects in DES Y3 is described in Sevilla-Noarbe et al. (2020).
- The Point-Spread Function (PSF) modeling algorithm and its validation tests are described in Jarvis et al. (2021).

<sup>1</sup> <https://www.darkenergysurvey.org/>

<sup>2</sup> <http://kids.strw.leidenuniv.nl/DR4/index.php>

<sup>3</sup> <https://www.naoj.org/Projects/HSC>

- A suite of image simulations, used to test the shape measurement pipeline and ultimately determine the shear calibration uncertainties is described in [MacCrann et al. \(2020\)](#).
- The METACALIBRATION shape catalog, and the tests that validate its science-readiness, are described in [Gatti, Sheldon et al. \(2021\)](#). This paper also discusses the (first layer) catalog-level blinding implemented in Y3.
- The characterization of the source redshift distribution, and the related systematic and statistical uncertainties, are detailed in five papers. Namely, [Myles, Alarcon et al. \(2020\)](#) and [Buchs, Davis et al. \(2019\)](#) present the baseline methodology for estimating wide-field redshift distributions using Self-Organizing Maps; [Gatti, Gianini et al. \(2020\)](#) outline an alternative method using cross correlations with spectroscopic galaxies; [Sánchez, Prat et al. \(2021\)](#) presents a complementary likelihood using small scale galaxy-galaxy lensing, improving constraints on redshifts and IA; finally [Cordero, Harrison et al. \(2021\)](#) validates our fiducial error parameterization using a more complete alternative based on distribution realizations. In addition to this, [Hartley, Choi et al. \(2020\)](#) and [Everett et al. \(2020\)](#) respectively describe the DES deep fields and the Balrog image simulations, both of which are crucial in testing and implementing the Y3 redshift methodology.
- The data covariance matrix is described in [Friedrich et al. \(2020\)](#). This paper also presents various validation tests based on DES Y3 simulations, and demonstrates its suitability for likelihood analyses.
- The numerical metrics used to assess tension between our DES results and external data sets are described in [Lemos et al. \(2020\)](#). That work considers a number of alternatives, and sets out the methodology used in this paper and [Dark Energy Survey Collaboration \(2021a\)](#).
- The simultaneous blinding of the multiple DES Y3 probes at the two-point correlation function level is described in [Muir et al. \(2020\)](#);
- [DeRose et al. \(2021a\)](#) presents a set of cosmological simulations which are used as an end-to-end validation of our analysis framework on mock  $N$ -body data.
- Finally, tests of the theoretical and numerical methods, as well as modeling assumptions for all  $3 \times 2$ pt analyses are described in [Krause et al. \(2021\)](#).

This paper is organized as follows: Sec. II describes the DES Y3 data, and the catalog construction and calibration. Sec. III describes the two-point measurements upon which our results are based, as well as the covariance estimation and blinding scheme. In Sec. IV, we describe the theoretical modeling of the cosmic shear two-point data vector. We demonstrate our model is robust to various forms of systematic

error, using simulated data, in Sec. V. Our baseline results and an exploration of the IA model complexity present in our data are then presented in Sec. VI. In Sec. VII we present a series of reanalyses, using slightly different modeling choices, in order to verify the robustness of our findings. The consistency of DES Y3 cosmic shear data with external probes such as other weak lensing surveys and the CMB is examined in Sec. VIII. Finally, Sec. IX summarizes our findings and discusses their significance in the context of the field.

## II. DES Y3 DATA & SAMPLE SELECTION

This section briefly describes the DES Y3 data, and defines the galaxy samples used in this paper. We also discuss a number of related topics, including calibrating selection biases.

### A. Data Collection & the GOLD Selection

DES has now completed its six-year campaign, covering a footprint of around  $5000 \text{ deg}^2$  to a depth of  $r \sim 24.4$ . The DES data were collected using the 570 megapixel Dark Energy Camera (DECam; [Flaugher et al. \(2015\)](#)), at the Blanco telescope at the Cerro Tololo Inter-American Observatory (CTIO), Chile, using five photometric filters *grizY*, which cover a region of the optical and near infrared spectrum between 0.40 and  $1.06 \mu\text{m}$ . DES SV, Y1 and Y3 cover sequentially larger fractions of the full Y6 footprint, with Y3, the data set used in this analysis, encompassing  $4143 \text{ deg}^2$  after masking, with the “Wide Survey” footprint covered with 4 overlapping images in each band (compared with the final survey depth of  $\sim 8$ ). The images undergo a series of reduction and pre-processing steps, including background subtraction ([Bernstein et al. 2018](#), [Eckert et al. 2020](#), [Morganson et al. 2018](#)), and masking out cosmic rays, satellite trails and bright stars. Object detection is performed on the *riz* coadd images using Source Extractor ([Bertin & Arnouts 1996](#)). For the detected galaxies, derived photometric measurements are generated using Multi-Object Fitting (MOF; [Drlica-Wagner et al. 2018](#)) to mitigate blending. The final Y3 selection with baseline masking is referred to as the GOLD catalog, and is described in detail in [Sevilla-Noarbe et al. \(2020\)](#).

### B. Shape Catalog & Image Simulations

The DES Y3 shape catalog is created using the METACALIBRATION algorithm ([Huff & Mandelbaum 2017](#), [Sheldon & Huff 2017](#)). The basic shape measurement entails fitting a single elliptical Gaussian to each detected galaxy. The fit is repeated on artificially sheared copies of the given galaxy, in order to construct a shear response matrix  $R_\gamma$  via a numerical derivative; a selection response  $R_S$  is also computed in a similar way. These multiplicative responses are the essence of METACALIBRATION. After quality cuts, the Y3 METACALIBRATION catalog contains over 100 million galaxies, with a

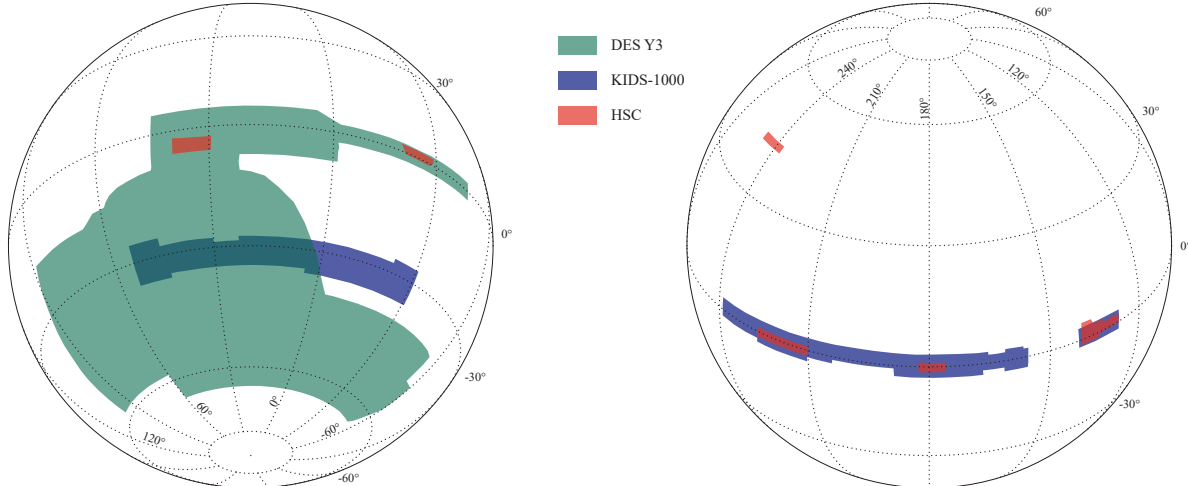


FIG. 1. The approximate footprints of Stage-III dark energy experiments: Dark Energy Survey Year 3 (DES Y3; green), Kilo-Degree Survey (KiDS-1000; blue) and first-year Hyper Suprime-Cam Subaru Strategic Program (HSC; red). The left and right panels show orthographic projections of the northern and southern sky respectively. The parallels and meridians show declination and right ascension. The different survey areas not only affect the final analysis choices, but also reflect the individual science strategies and the complementarity of Stage-III surveys.

mean redshift of  $z = 0.63$  and a weighted number density<sup>4</sup>  $n_{\text{eff}} = 5.59 \text{ arcmin}^{-2}$ ; for discussion of the cuts and why they are necessary, see [Gatti, Sheldon et al. \(2021\)](#).

Although METACALIBRATION greatly reduces the biases inherent to shear estimation, the process is not perfect. We must still rely on image simulations for validation and for deriving priors on the residual biases (predominantly due to blending, and its impact on the redshift distribution). These simulations, and the conclusions we draw from them for Y3, are discussed in [MacCrann et al. \(2020\)](#). In addition to tests using simulations, the catalogs are subject to a number of null tests, applied directly to the data. Using both pseudo- $C_\ell$ s and COSEBIs ([Schneider et al. 2010](#)), we find no evidence for non-zero B-modes in Y3.

### C. Photometric Redshift Calibration

We estimate and calibrate the redshift distributions of our source sample with a combination of three different methods. Our base methodology is known as *Self-Organizing Map*  $p(z)$  (SOMPZ; [Myles, Alarcon et al. 2020](#)). The most important aspect of this methodology is that knowledge from precise redshifts (from spectroscopic samples) and higher-quality photometric data (from DES deep fields [Hartley, Choi et al. 2020](#))

informs the bulk of the DES observations (the wide fields), essentially acting as a Bayesian prior. The connection between the deep and wide field data is determined empirically using an image simulation framework known as Balrog ([Everett et al. 2020](#)).

Additionally, clustering redshifts (WZ; [Gatti, Giannini et al. 2020](#)) employ cross-correlations of galaxy densities to improve redshift constraints, and shear ratios ([Sánchez, Prat et al. 2021](#)) help to constrain redshifts (and also intrinsic alignment parameters), utilizing galaxy-shear correlation functions at small scales. While SOMPZ and WZ are applied upstream to generate and select  $n(z)$  estimates, the shear-ratio information, on the other hand, is incorporated at the point of evaluation of cosmological likelihoods (see Sec. IV F). Details of each of these methods in the context of Y3 can be found in [Myles, Alarcon et al. \(2020\)](#), and robustness tests of redshift distributions in the context of cosmic shear are presented in [Amon et al. \(2021\)](#).

## III. COSMIC SHEAR MEASUREMENT

In this section we present the measured real-space cosmic shear two-point correlations ( $\xi_{\pm}(\theta)$ , see Eq. 11), which form the basis of our results and are shown in Fig. 2. Defining the signal-to-noise of our measurement as

$$S/N \equiv \frac{\xi_{\pm}^{\text{data}}(\theta)^T \mathbf{C}^{-1} \xi_{\pm}^{\text{model}}(\theta)}{\sqrt{\xi_{\pm}^{\text{model}}(\theta)^T \mathbf{C}^{-1} \xi_{\pm}^{\text{model}}(\theta)}}, \quad (1)$$

<sup>4</sup> The effective number density here is as defined by [Heymans et al. \(2012\)](#). The equivalent value using [Chang et al. \(2013\)](#)'s definition is  $5.32 \text{ arcmin}^{-2}$  (see [Gatti, Sheldon et al. 2021](#) for details.)

where  $\mathbf{C}$  is the data covariance matrix, the  $S/N$  of the cosmic shear detection in DES Y3 after scale cuts is 27. For the fiducial  $\Lambda$ CDM model, our chi-square at the maximum posterior is  $\chi^2 = 237.7$ , with 222 effective degrees-of-freedom (d.o.f.), which gives us a  $p$ -value of 0.22 (see Sec. VIA). We define these quantities in more detail in Sec. IV.

### A. Tomography

We define a set of four broad redshift bins for our source sample in the nominal range  $0 < z < 3$ , with actual number densities being fairly small above  $z \gtrsim 1.5$ . These are constructed by iteratively adjusting the redshift bin edges, such that they each yield approximately the same number of source galaxies. The Y3 SOMPZ methodology (see Section IIC) makes use of Balrog, which artificially inserts COSMOS galaxies into DES images. The artificial galaxies are assigned to cells in both the wide- and deep-field Self-Organizing Maps (SOMs), which allows one to map between the two, and so assign DES wide-field galaxies to bins (see Myles, Alarcon et al. 2020, Sec. 4.3).

The redshift distributions computed in this way, which feed into our modeling in the next section, are shown in Fig. 3. The galaxy number densities are 1.476, 1.479, 1.484 and 1.461 per square arcminute respectively in these four redshift bins.

### B. Two-Point Estimator & Measurement

The spin-2 shear field can be expressed in terms of a real and an imaginary component,  $\gamma = \gamma_1 + i\gamma_2$ . There are two possible shear two-point functions that preserve parity invariance, and a ‘‘natural’’ convention for them is:  $\xi_+ \equiv \langle \gamma\gamma^* \rangle$  and  $\xi_- \equiv \langle \gamma\gamma \rangle$  (Schneider & Lombardi 2003), where the angle brackets denote averaging over galaxy pairs. In terms of tangential ( $t$ ) and cross ( $\times$ ) components defined along the line that connects each pair of galaxies  $a, b$ , we have:

$$\xi_{\pm} = \langle \gamma_{t,a}\gamma_{t,b} \rangle_{ab} \pm \langle \gamma_{\times,a}\gamma_{\times,b} \rangle_{ab}. \quad (2)$$

In practice we do not have direct access to the shear field, but rather estimate it via per-galaxy ellipticities (although see Bernstein et al. 2016 for an alternative approach). Correlating galaxies in a pair of redshift bins ( $i, j$ ) we define,

$$\xi_{\pm}^{ij}(\theta) = \frac{\sum_{ab} w_a w_b \left( \hat{e}_{t,a}^i \hat{e}_{t,b}^j \pm \hat{e}_{\times,a}^i \hat{e}_{\times,b}^j \right)}{\sum_{ab} w_a w_b R_a R_b}, \quad (3)$$

with inverse variance weighting  $w^5$  (unlike in Y1, where such weighting was not included) and response factors  $R$  that account for shear and selection biases (see Gatti, Sheldon et al.

2021 for details), and where the sums run over pairs of galaxies  $a, b$ , for which the angular separation falls within the range  $|\theta - \Delta\theta|$  and  $|\theta + \Delta\theta|$ . Both  $\xi_+$  and  $\xi_-$  are measured using twenty log-spaced  $\theta$  bins between 2.5 and 250 arcminutes, with  $i, j \in (1, 2, 3, 4)$ . As discussed later, not all of the twenty angular bins are utilized in our likelihood analysis. We also assume the response matrix is diagonal and that the selection part is scale independent. The ellipticities that enter Eq. (3) are corrected for residual mean shear, such that  $\hat{e}_k^i \equiv e_k^i - \langle e_k \rangle_i$  for components  $k \in (1, 2)$  and redshift bin  $i$ , again following the Y1 methodology (Troxel et al. 2018). We show the resulting two-point functions, which are measured using TreeCorr<sup>6</sup> (Jarvis et al. 2004), in Fig. 2, alongside best fitting theory predictions.

### C. Data Covariance Matrix

We model the statistical uncertainties in our combined measurements of  $\xi_{\pm}$  as a multivariate Gaussian distribution. The disconnected 4-point function part of the covariance matrix of that data vector (the Gaussian covariance part) is described in Friedrich et al. 2020 and includes analytic treatment of bin averaging and sky curvature. We also verify in that paper that expected fluctuations in  $\Delta\chi^2$  between the measurement and our maximum posterior model do not significantly impact on our estimates of cosmological parameters. Our modeling of the connected 4-point function part of the covariance matrix and the contribution from super-sample covariance uses the public CosmoCov<sup>7</sup> (Fang et al. 2020) code, which is based on the CosmoLike framework (Krause & Eifler 2017).

We use the RMS per-component shape dispersion  $\sigma_e$  and effective number densities  $n_{\text{eff}}$  specified in Table 1 of Amon et al. 2021 to calculate the shape-noise contribution to the covariance, and additionally account for survey geometry effects. We follow previous cosmic shear analyses in using a covariance matrix that assumes a baseline cosmology (see Hikage et al. 2019 for a different approach). That is, we assume a fiducial set of input parameters for the initial covariance matrix and run cosmological chains using this first guess. The covariance is then recomputed at the best fit from this first iteration, and the final chains are run. We find this update to have negligible effects on the cosmic shear constraints presented in this paper.

### D. Blinding

We implement a three-stage blinding strategy, performing transformations to the catalog, data vector, and parameters in order to obscure the cosmological results of the analysis. By disconnecting the people carrying out the analysis from the impact their various choices are having on the eventual cosmological results, the aim is to avoid unconscious biases,

<sup>5</sup> Although referred to as such, the catalog weights only approximate inverse variance weighting. See Gatti, Sheldon et al. (2021), Sec 4.3 for details.

<sup>6</sup> <https://github.com/rmjarvis/TreeCorr>

<sup>7</sup> <https://github.com/CosmoLike/CosmoCov>

either towards or away from previous results in the literature. Although the approaches differ somewhat, all of the major cosmic shear collaborations have adopted a similar philosophy regarding the necessity of blinding (Asgari et al. 2021, Dark Energy Survey Collaboration 2016, Hikage et al. 2019, Hildebrandt et al. 2020, Troxel et al. 2018).

The first level of blinding follows a similar method to that used in Y1 (Zuntz et al. 2018), and is discussed in Gatti, Sheldon et al. (2021) (their Sec. 2.3). In short, the process involves a transformation of the shear catalog, where galaxy shapes are scaled by a random multiplicative factor. The second level is a transformation of the data vector using the method described in Muir et al. (2020). We compute model predictions at two sets of input parameters: an arbitrary reference cosmology  $\Theta_{\text{ref}}$  and a shifted cosmology  $\Theta_{\text{ref}} + \Delta\Theta$ , where  $\Delta\Theta$  is drawn randomly in  $w$ CDM parameter space. The difference between these model predictions is then applied to the measured  $\xi_{\pm}$  data vector prior to its analysis. The final stage of blinding is at the parameter level, and entails obscuring the axes of contour plots (effectively equivalent to shifting contours randomly in parameter space, preserving constraining power but making external consistency testing impossible). A detailed checklist of the tests that must be fulfilled before each stage of blinding can be removed can be found in Dark Energy Survey Collaboration (2021a).

From a modeling perspective, passing the tests we describe in Sec. V and the further tests on synthetic data described in DeRose et al. (2021a) and Krause et al. (2021) fulfills our unblinding requirements. A set of internal consistency tests for cosmic shear must also be passed and are described in Amon et al. (2021).

#### IV. FIDUCIAL MODEL AND ANALYSIS CHOICES

A predictive physical model for cosmic shear has a number of requirements; first, any systematic deviations or effects omitted from the model must be comfortably subdominant to uncertainties on the data; and second, the implementation must be numerically stable at all points within the prior volume. We also aim for redundancy, and implement the full pipeline in two independent codes: CosmoSIS<sup>8</sup> (Zuntz et al. 2015) and CosmoLike (Krause & Eifler 2017), which are verified to be in agreement to within a negligible  $\Delta\chi^2$  (Krause et al. 2021).

In this section, we outline our baseline model for  $\xi_{\pm}^{ij}(\theta)$  and discuss how it meets the above criteria. We subsequently show the cosmological constraints from these analysis choices in Sec. VI. To test the robustness of this baseline model, we later relax its main approximations and assumptions and show variations of analysis choices in Sec. VII.

#### A. Sampling and Parameter Inference

For all parameter inference presented in this paper, we assume the likelihood of the data given the model  $M$  with parameters  $\mathbf{p}$  to be a multivariate Gaussian:

$$\ln \mathcal{L}(\hat{\mathbf{D}}|\mathbf{p}, M) = -\frac{1}{2}\chi^2 + \text{const.}, \quad (4)$$

$$\chi^2 = \left(\hat{\mathbf{D}} - \mathbf{T}_M(\mathbf{p})\right)^T \mathbf{C}^{-1} \left(\hat{\mathbf{D}} - \mathbf{T}_M(\mathbf{p})\right) \quad (5)$$

where  $\mathbf{C}$  is the data covariance matrix and  $\mathbf{T}_M(\mathbf{p})$  is the theory prediction vector for a data vector  $\hat{\mathbf{D}}$ , a concatenated version of all elements of the tomographic cosmic shear data (with length  $N_D = N_{\theta}N_z(N_z + 1)$ , where  $N_{\theta}$  is the number of angular bins included in each correlation function after scale cuts ( $N_{\theta}$  varies depending on the redshift bins, and equals 20 before cuts) and  $N_z = 4$  is the number of broad redshift bins). Since we are incorporating shear ratios at the inference level (see Sec. IV F), the final likelihood used in our analyses is the sum of two parts,  $\ln \mathcal{L} = \ln \mathcal{L}_{2\text{pt}} + \ln \mathcal{L}_{\text{SR}}$ , which are assumed to be independent (Sánchez, Prat et al. 2021).

As we aim to perform a Bayesian analysis, the *a posteriori* knowledge of the parameters given the observed data, denoted by  $\mathcal{P}(\mathbf{p}|\hat{\mathbf{D}}, M)$ , depends not only on the likelihood but also on prior  $\Pi(\mathbf{p}|M)$ . These pieces are related via Bayes' theorem:

$$\mathcal{P}(\mathbf{p}|\hat{\mathbf{D}}, M) = \frac{\mathcal{L}(\hat{\mathbf{D}}|\mathbf{p}, M)\Pi(\mathbf{p}|M)}{P(\hat{\mathbf{D}}|M)}, \quad (6)$$

where  $P(\hat{\mathbf{D}}|M)$  is the so-called evidence of the data. Sampling of the posterior is carried out using POLYCHORD (Handley et al. 2015; with 500 live points and tolerance 0.01). These settings have been tested to demonstrate the accuracy of the posteriors and Bayesian evidence estimates. Although sampling gives a rather noisy estimate of the best-fit point in the full parameter space, the *Maximum a Posteriori* (MAP) quoted in Sec. VI A, we verify that standard POLYCHORD outputs, in practice, offer a reasonable estimate of that point when compared to a MAP optimizer<sup>9</sup>. Throughout this paper, we report parameter constraints using the MAP value and 1D marginalized summary statistics in the form:

$$\text{Parameter} = \text{1D mean}^{\text{+upper 34\% bound}}_{\text{-lower 34\% bound}} \text{ (MAP value)}.$$

The ratio of evidences is a well-defined quantity for model-testing within a single data set to indicate a preference for one modeling choice ( $M_1$ , with parameters  $\mathbf{p}_1$ ) over another ( $M_2$ ,  $\mathbf{p}_2$ ):

$$R_{M_1/M_2} = \frac{P(\hat{\mathbf{D}}|M_1)}{P(\hat{\mathbf{D}}|M_2)} = \frac{\int d\mathbf{p} \mathcal{L}(\hat{\mathbf{D}}|\mathbf{p}_1, M_1)P(\mathbf{p}_1|M_1)}{\int d\mathbf{p} \mathcal{L}(\hat{\mathbf{D}}|\mathbf{p}_2, M_2)P(\mathbf{p}_2|M_2)}. \quad (7)$$

<sup>8</sup> <https://bitbucket.org/joezuntz/cosmosis>

<sup>9</sup> We run the MAXLIKE sampler in Posterior mode for this test (<https://bitbucket.org/joezuntz/cosmosis/wiki/samplers/maxlike>)



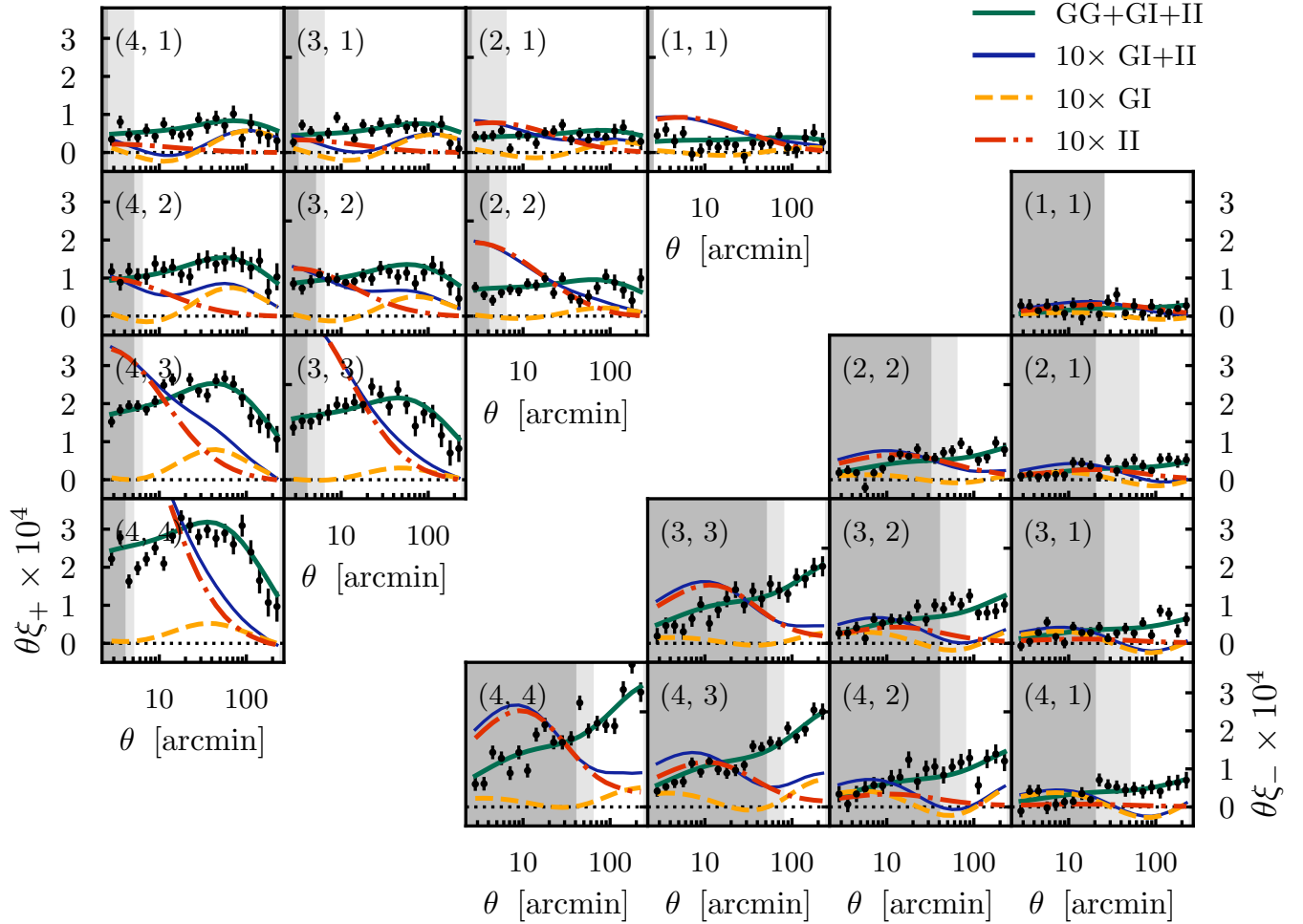


FIG. 2. Cosmic shear two-point correlation measurements from DES Y3. We show here  $\xi_+$  and  $\xi_-$  (black data points, upper left and lower right halves respectively), with the different panels showing different combinations of redshift bins; in all cases the error bars come from our fiducial analytic covariance matrix. The lighter grey bands represent scales removed from our fiducial analysis, while the darker are the equivalent for the  $\Lambda$ CDM Optimized analysis. Also shown are the best-fit theory curve in  $\Lambda$ CDM (solid green) and the intrinsic alignment contributions to the signal: GI (dashed yellow), II (dot-dashed red) and GI+II (solid blue). For clarity, we multiply the IA contributions by a factor of 10, and in most bins the total IA signal is  $\sim 1\%$  of GG+GI+II. The detection significance of the cosmic shear signal after fiducial scale cuts is 27. The  $\chi^2$  per effective d.o.f of the  $\Lambda$ CDM model is  $237.7/222 = 1.07$  (a  $p$ -value of 0.22).

The evidence ratio has the advantage of naturally penalizing models of excessive parameter space volume, but needs to be interpreted using e.g. the Jeffreys scale (Jeffreys 1935), which somewhat arbitrarily differentiates between “strong” and “weak” model preferences. Our main use of evidence ratios is to help assessing model preference in the context of IA complexity in Sec. VIC2.

### B. Modeling Cosmic Shear

The two-point cosmic shear correlations  $\xi_{\pm}^{ij}(\theta)$  are related to the nonlinear matter power spectrum (and thus to the growth and evolution of structure). The key quantity that dictates how much a galaxy on a particular line of sight is distorted, is known as the *convergence*  $\kappa$ . That is, the weighted mass overdensity  $\delta$ ,

integrated along the line-of-sight to the distance of the source  $\chi_s$ :

$$\kappa(\theta) = \int_0^{\chi_s} d\chi W(\chi)\delta(\theta, \chi). \quad (8)$$

The weight for a particular lens plane, quantified below in Eq. (10), is sensitive to the relative distances of the source and the lens; it is via this geometrical term that cosmic shear probes the expansion history of the Universe. Fitting all the auto- and cross-redshift bin correlations simultaneously significantly improves the cosmological constraining power (Hu 1999), both because it helps to untangle the signal at different epochs, and because it (partially) breaks the degeneracy with intrinsic alignments (see also Sec. IVD).

Under the Limber approximation (Limber 1953, LoVerde & Afshordi 2008), the 2D convergence power spectrum in

tomographic bins  $i$  and  $j$ ,  $C_\kappa^{ij}(\ell)$  is related to the full 3D matter power spectrum as:

$$C_\kappa^{ij}(\ell) = \int_0^{\chi(z_{\max})} d\chi \frac{W^i(\chi)W^j(\chi)}{\chi^2} P_\delta \left( \frac{\ell+1/2}{\chi}, z(\chi) \right), \quad (9)$$

where  $P_\delta$  is the nonlinear matter power spectrum and the lensing efficiency kernels are given by

$$W^i(\chi) = \frac{3H_0^2\Omega_m}{2c^2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_H} d\chi' n^i(z(\chi')) \frac{dz}{d\chi'} \frac{\chi' - \chi}{\chi'}. \quad (10)$$

The source galaxy redshift distribution  $n^i(z)$  here is normalized to unity. Clearly, the amplitude of  $C_\kappa$  responds directly to  $\sigma_8^2$ , and to  $\Omega_m$  via the power spectrum, and to  $\Omega_m h^2$  via the lensing kernel, which gives rise to a characteristic banana-shaped degeneracy in  $\sigma_8 \times \Omega_m$ . The combination of parameters most strongly constrained by cosmic shear is a derived parameter, commonly referred to as  $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ .

By decomposing  $\kappa$  into E- and B-mode components (Crittenden et al. 2002, Schneider et al. 2002) and in a full-sky formalism, one can express the angular two point shear correlations as:

$$\xi_\pm^{ij}(\theta) = \sum_\ell \frac{2\ell+1}{2\pi\ell^2(\ell+1)^2} [G_{\ell,2}^+(\cos\theta) \pm G_{\ell,2}^-(\cos\theta)] \times [C_{EE}^{ij}(\ell) \pm C_{BB}^{ij}(\ell)], \quad (11)$$

where the functions  $G_\ell^\pm(x)$  are computed from Legendre polynomials  $P_\ell(x)$  and averaged over angular bins (see Krause et al. (2021) Eqs. 19 and 20). It is also worth bearing in mind that, in practice, the angular spectra in Eq. (11) are not in fact pure cosmological convergence spectra  $C_\kappa$ , but rather shear spectra  $C_\gamma$ , which include contributions from intrinsic alignments (see Sec. IV D), and additional higher order terms are explored later.

Since the lensing kernel in Eq. (10) acts as a redshift filter, which modulates sensitivity to  $P_\delta$  in Eq. (9), it is informative to show its redshift dependence; we do so in the lower panel of Fig. 3, for the fiducial Y3 redshift distributions. As shown there, the DES Y3 cosmic shear signal is sensitive to a relatively broad range, with kernels peaking between approximately  $z = 0.1$  and  $z = 0.5$ .

Similarly, the polynomials in Eq. (11) mix together a range of physical distances into any given angular scale. We can elucidate this by writing  $\xi_\pm^{ij}$  as an integral over  $\ln k$  (e.g. Tegmark & Zaldarriaga 2002) to obtain :

$$\xi_\pm^{ij}(\theta) = \int_{-\infty}^{+\infty} d \ln k \mathcal{P}(k) \quad (12)$$

where

$$\mathcal{P}(k) = \frac{k}{2\pi} \sum_\ell [G_{\ell,2}^+(\cos\theta) \pm G_{\ell,2}^-(\cos\theta)] \times W^i(\chi)W^j(\chi)P_\delta(k, z(\chi)), \quad (13)$$

with  $\chi = (\ell + 1/2)/k$ . We show  $d \ln \xi_\pm(\theta)/d \ln k = \mathcal{P}(k)/\xi_\pm(\theta)$  for representative  $\theta$  scales in Fig. 4. This shows

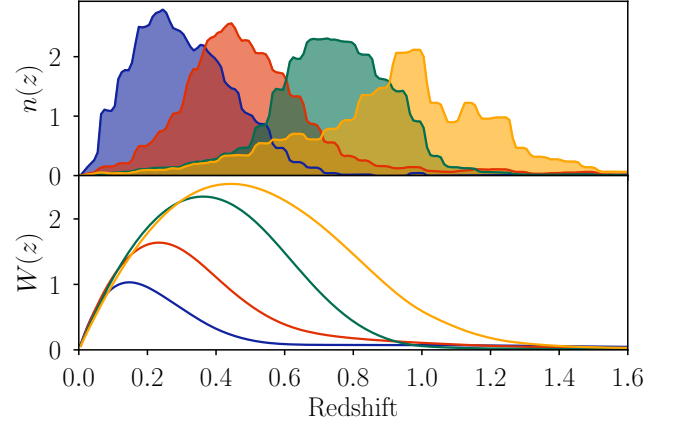


FIG. 3. The estimated redshift distributions and lensing kernels for the fiducial source galaxy sample used in this work. Most of the sensitivity of the DES Y3 cosmic shear signal to large scale structure is in the range between  $z = 0.1$  and  $z = 0.5$ , where individual kernels peak. Each distribution is independently normalized over the redshift range  $z = 0 - 3$ . The total effective number density (Heymans et al. 2012) of sources is  $n_{\text{eff}} = 5.59$  galaxies per square arcminute and is divided almost equally into the 4 redshift bins.

the sensitivity of our cosmic shear signal at a given angular scale to modes of the matter power spectrum. Our scale cuts, defined in Sec. IV G, eventually remove most of the sensitivity to  $k > 1 \text{ hMpc}^{-1}$ .

The remainder of this Sec. motivates the ingredients introduced in Eq. (9) to (11) such as the non-linear power spectrum prescription and the set of scales for which the signal is not significantly contaminated by unmodeled physics.

### C. Nonlinear Power Spectrum

On the largest of physical scales, growth is linear and well described by a purely linear matter power spectrum  $P_\delta^{\text{lin}}(k)$ . To evaluate  $P_\delta^{\text{lin}}$  we use the Boltzmann code CAMB<sup>10</sup> (Lewis et al. 2000), as implemented in CosmoSIS. On smaller scales, however, this is not true, and one also needs a model for non-linear growth. Our fiducial model for the non-linear matter power spectrum  $P_\delta(k)$  is the HALOFIT functional prescription (Smith et al. 2003, Takahashi et al. 2012). We have made scale cuts to remove the parts of the data vector affected by baryonic effects, as described in Sec. IV G; this largely removes the sensitivity to wavenumbers  $k > 1 \text{ hMpc}^{-1}$ . For  $k < 1 \text{ hMpc}^{-1}$ , Takahashi et al. (2012) reports an uncertainty on the HALOFIT model of 5%. In Krause et al. (2021) we demonstrate, by substituting HALOFIT for HMCODE<sup>11</sup> (Mead et al. 2015), the EUCLID EMULATOR (Euclid Collaboration 2019), or the MIRATITAN EMULATOR (Lawrence et al. 2017) that for cosmic shear alone we are insensitive to this choice. In the context of the Y3

<sup>10</sup> <http://camb.info>

<sup>11</sup> <https://github.com/alexander-mead/HMcode>

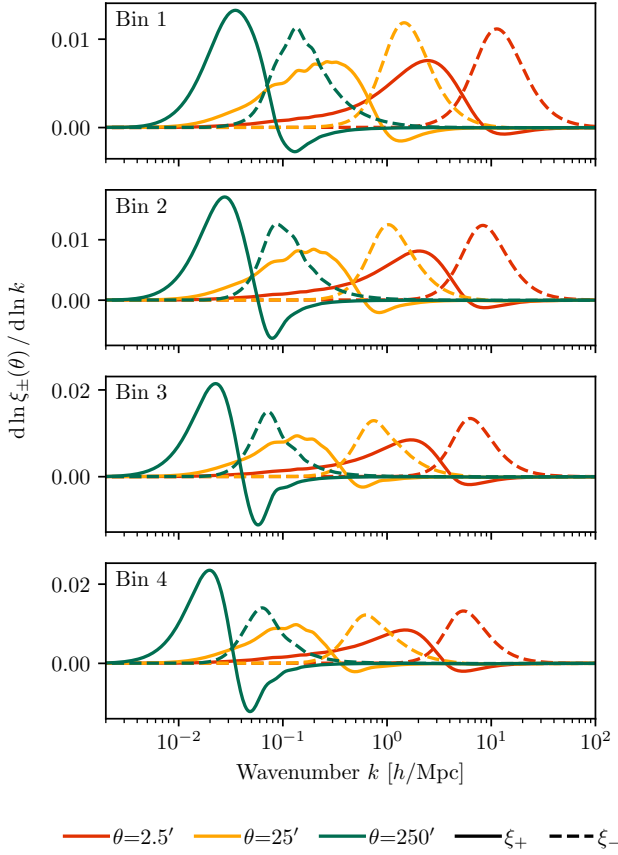


FIG. 4. Window functions of  $\xi_+$  (solid curves) and  $\xi_-$  (dashed curves) over  $k$ -wavenumbers of the matter power spectrum at representative angular separations. Notice that our smallest angular scales (after cuts) in  $\xi_+$  and  $\xi_-$  are around 2.5 arcminutes and 30 arcminutes respectively, which means that only a relatively small contribution to the full signal comes from wavenumbers above  $k \sim 1$  h/Mpc.

$3 \times 2$ pt analysis, the distinction between these three models is more nuanced, and we refer the reader to Krause et al. (2021) for a full justification of the use of HALOFIT.

#### D. Intrinsic Alignments

Galaxies are not idealized tracers of the underlying matter field, but rather astrophysical bodies, which are subject to local interactions. To account for this added complexity, the observed shape of a galaxy can be decomposed into two parts, the shear induced by gravitational lensing (G) and the intrinsic shape (I) induced by the local environment:  $\gamma = \gamma^G + \gamma^I$ . In this section, we consider only the correlated intrinsic component, and not the intrinsic “shape noise”, which contributes to the covariance but not the signal.

The term *intrinsic alignments* covers two contributions from environmental interactions: (a) intrinsic shape - intrinsic shape correlations between galaxies that are physically close to each other, and (b) shear-intrinsic correlations between galaxies on

neighbouring lines of sight. Known as II and GI contributions respectively, and contributing on similar angular scales to the cosmological lensing signal, these terms constitute a significant systematic in weak lensing analyses. Including IA contributions, the observed E-mode angular power spectrum is written

$$C_{\gamma,EE}^{ij}(\ell) = C_{GG}^{ij}(\ell) + C_{GI}^{ij}(\ell) + C_{IG}^{ij}(\ell) + C_{II,EE}^{ij}(\ell). \quad (14)$$

Nonlinear models of IA, as discussed below, can also produce a non-zero B-mode power spectrum:

$$C_{\gamma,BB}^{ij}(\ell) = C_{II,BB}^{ij}(\ell). \quad (15)$$

Assuming the Limber approximation as before, the two IA  $C(\ell)$ s are given by:

$$C_{GI}^{ij}(\ell) = \int_0^{\chi_H} d\chi \frac{W^i(\chi)n^j(\chi)}{\chi^2} P_{GI}\left(\frac{\ell+1/2}{\chi}, z(\chi)\right), \quad (16)$$

and

$$C_{II}^{ij}(\ell) = \int_0^{\chi_H} d\chi \frac{n^i(\chi)n^j(\chi)}{\chi^2} P_{II}\left(\frac{\ell+1/2}{\chi}, z(\chi)\right), \quad (17)$$

These expressions are generic, and are valid regardless of which model is used to predict  $P_{GI}$  and  $P_{II}$  (see the following subsections).

##### 1. IA and the tidal field

It is typically assumed that the correlated component of galaxy shapes is determined by the large-scale cosmological tidal field. The simplest relationship, which should dominate on large scales and for central galaxies, involves the “tidal alignment” of galaxy shapes, producing a linear dependence (Catelan et al. 2001, Hirata & Seljak 2004). In this case, one can relate the intrinsic shape component to the gravitational potential at the assumed time of galaxy formation  $\phi_*$ :

$$(\gamma_1^I, \gamma_2^I) = A_1(z) \left( \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}, 2 \frac{\partial^2}{\partial x \partial y} \right) \phi_*, \quad (18)$$

where the proportionality factor  $A_1(z)$  captures the response of intrinsic shape to the tidal field. More complex alignment processes, including “tidal torquing,” relevant for determining the angular momentum of spiral galaxies, are captured in a nonlinear perturbative framework, which we refer to as “TATT” (Tidal Alignment and Tidal Torquing; Blazek et al. 2019). In this more general model, we use nonlinear cosmological perturbation theory to express the intrinsic galaxy shape field, measured at the location of source galaxies (Blazek et al. 2015), as an expansion in the matter density field  $\delta$  and tidal field  $s_{ij}$ :

$$\tilde{\gamma}_{ij}^{IA} = A_1 s_{ij} + A_1 \delta s_{ij} + A_2 \sum_k s_{ik} s_{kj} + \dots, \quad (19)$$

where  $s_{ij}$  is the gravitational tidal field, which at any given position  $\mathbf{x}$  is a  $3 \times 3$  tensor (see [Catelan & Porciani 2001](#) for a formal definition).

Although the terms in the model can be associated with physical mechanisms, they can also be viewed as effective contributions to intrinsic shape correlations from small-scale physics. See also [Schmitz et al. \(2018\)](#), [Tugendhat & Schäfer \(2018\)](#), [Vlah et al. \(2020\)](#) for further discussion of the perturbative approach and [Fortuna et al. \(2021\)](#) for a halo model treatment of IA.

## 2. Model implementation: NLA and TATT

Within the TATT framework, three parameters capture the relevant responses to the large-scale tidal fields (see [Blazek et al. 2019](#) for more details):  $A_1$ ,  $A_2$ , and  $A_{1\delta}$ , corresponding respectively to a linear response to the tidal field (tidal alignment), a quadratic response (tidal torquing), and a response to the product of the density and tidal fields. To date, the most frequently used intrinsic alignment model in the literature is known as the Nonlinear Alignment Model (NLA; [Bridle & King 2007](#), [Hirata et al. 2007](#)), an empirically-based modification of the Linear Alignment (LA) model of [Catelan et al. \(2001\)](#) and [Hirata & Seljak \(2004\)](#), in which the fully nonlinear tidal field is used to calculate the tidal alignment term. Within the ‘‘TATT’’ framework, the NLA model corresponds to only  $A_1$  being non-zero in Eq. 19. The GI and II power spectra then have the same shape as the nonlinear matter power spectrum, but are modulated by  $A_1(z)$ :

$$P_{\text{GI}}(k, z) = A_1(z)P_\delta(k, z), \quad P_{\text{II}}(k, z) = A_1^2(z)P_\delta(k, z). \quad (20)$$

Note that the nonlinear power spectrum and the IA amplitudes are functions of redshift. In the following, the  $z$  dependence of the IA amplitudes and various  $k$  dependent terms is left implicit. More generally, in the TATT model, the GI and II power spectra are constructed with the relevant correlations of tidal and density fields:

$$P_{\text{GI}}(k) = A_1 P_\delta(k) + A_{1\delta} P_{0|0E}(k) + A_2 P_{0|E2}(k), \quad (21)$$

$$P_{\text{II,EE}}(k) = A_1^2 P_\delta(k) + 2A_1 A_{1\delta} P_{0|0E}(k) + A_{1\delta}^2 P_{0E|0E}(k) \\ + A_2^2 P_{E2|E2}(k) + 2A_1 A_2 P_{0|E2}(k) \\ + 2A_{1\delta} A_2 P_{0E|E2}(k), \quad (22)$$

$$P_{\text{II,BB}}(k) = A_{1\delta}^2 P_{0B|0B}(k) + A_2^2 P_{B2|B2}(k) + 2A_{1\delta} A_2 P_{0B|B2}(k). \quad (23)$$

In this work, these  $k$ -dependent terms are evaluated using FAST-PT ([Fang et al. 2017](#), [McEwen et al. 2016](#)), as implemented in CosmoSIS. The model, including the full expressions for these power spectra is set out in some depth in [Blazek et al. \(2019\)](#) (see their Eqs. 37–39 and appendix A), and we refer the reader to that paper for technical details. The  $k$ -dependent contributions are modulated by the redshift-dependent amplitudes  $A_1$ ,  $A_2$ , and  $A_{1\delta}$ . We define the first two with the following convention:

$$A_1(z) = -a_1 \bar{C}_1 \frac{\rho_{\text{crit}} \Omega_m}{D(z)} \left( \frac{1+z}{1+z_0} \right)^{\eta_1}, \quad (24)$$

$$A_2(z) = 5a_2 \bar{C}_1 \frac{\rho_{\text{crit}} \Omega_m}{D^2(z)} \left( \frac{1+z}{1+z_0} \right)^{\eta_2}, \quad (25)$$

where  $D(z)$  is the linear growth factor,  $\rho_{\text{crit}}$  is the critical density and  $\bar{C}_1$  is a normalisation constant, by convention fixed at  $\bar{C}_1 = 5 \times 10^{-14} M_\odot h^{-2} \text{Mpc}^2$ , obtained from SuperCOSMOS (see [Brown et al. 2002](#)). The leading factor of 5 in Eq. (25) is included to account for the difference in the windowed variance produced by the TA and TT power spectra. With this factor included, the TA and TT contributions to  $P_{\text{II}}$  at  $z = 0$ , averaged over this window, should be roughly equal if  $a_1 = a_2$ , aiding in the interpretation of the best fitting values. Note that this is a matter of convention only, and does not affect our final cosmological results. The denominator  $z_0$  is a pivot redshift, which we fix to the value 0.62<sup>12</sup>. The dimensionless amplitudes ( $a_1, a_2$ ) and power law indices ( $\eta_1, \eta_2$ ) are free parameters in this model.

As mentioned above, the model also includes the  $A_{1\delta}$  contribution, corresponding to the product of the density and tidal fields. This term was originally motivated by the modulation of the IA signal due to the galaxy density weighting (i.e. the fact that the shape field is preferentially sampled in overdense regions [Blazek et al. 2015](#)). In this case, within the TATT model, we have

$$A_{1\delta} = b_{\text{TA}} A_1, \quad (26)$$

where  $b_{\text{TA}}$  is the linear bias of source galaxies contributing to the tidal alignment signal. In our baseline analysis, rather than fixing  $b_{\text{TA}}$  to this bias value, we sample over it with a wide prior, allowing the  $A_{1\delta}$  contribution to capture a broader range of nonlinear alignment contributions. We note that this is a departure from previous studies to have used this model ([Blazek et al. 2019](#), [Samuroff et al. 2019](#), [Troxel et al. 2018](#)), all of which held  $b_{\text{TA}} = 1$  fixed. The motivation for this change is set out in Sec. V. As can be seen from Eq. (21) - (23), in the limit  $a_2, b_{\text{TA}} \rightarrow 0$ , the TATT model reduces to the NLA model. It is thus useful to think of NLA as a sub-space of the more complete TATT model, rather than a distinct, alternative model. Given the sensitivity of IAs to the details of the galaxy selection, and in the absence of informative priors, we choose to marginalize over all five IA parameters ( $a_1, a_2, \eta_1, \eta_2, b_{\text{TA}}$ ), governing the amplitude and redshift dependence of the IA terms, with wide flat priors (see Sec. IV H). While a redshift evolution in the form of a power law, captured by the index  $\eta_i$ , is a common assumption, the  $A_i(z)$  coefficients could, in theory, have a more complicated redshift dependence. We seek to test the impact of this assumption by rerunning our analysis with

<sup>12</sup> The value was chosen in DES Y1 to be approximately equal to the mean source redshift. We choose to maintain that value to allow for an easier comparison of the IA amplitudes with those results.

a more flexible parameterization, whereby the IA amplitude  $A_{1,i}$  in each redshift bin is allowed to vary independently. The results of this exercise can be found in Sec. VII B.

It is finally worth remarking that the TATT model predicts a non-zero B-mode power spectrum  $P_{\Pi, BB}$ . This extra signal component is incorporated into our modeling, and propagated into  $\xi_{\pm}$  via Eq. (11).  $P_{\Pi, BB}$  is expected to be small since the testing carried out in Gatti, Sheldon et al. 2021 points to no statistically significant detection of B-modes in the DES Y3 shape catalog.

A visual representation of the  $\Pi$  and GI signals, as predicted at the best-fitting point in parameter space from our fiducial analysis, can be found in Fig. 2.

### E. Modeling Nuisance Parameters: Shear Bias & photo- $z$ Error

In addition to the parameters associated with the cosmological and IA models, there are a number of additional parameters included in our fiducial analysis, which are intended to absorb known sources of uncertainty. First of all, we follow the majority of previous cosmic shear analyses, and parameterize errors in the redshift distributions as uniform shifts in their mean, which transform the assumed distribution in a given bin  $i$  as:

$$n^i(z) \rightarrow n^i(z - \Delta z^i). \quad (27)$$

This leaves us with four nuisance parameters ( $\Delta z^{1-4}$ ), which are marginalized with informative priors (see Sec. IV H). This simple marginalization scheme was validated with a more complete method, known as HYPERRANK Cordero, Harrison et al. (2021), which uses a large ensemble of possible redshift distributions (typically of order 1000) generated using the SOMPS pipeline (Myles, Alarcon et al. 2020). The basic aim of this method is to capture the range of plausible variations in the full shape of the distributions (not only their mean), and also the correlations between bins. Although in our particular setup the simpler parameterization was found to be sufficient, the tests in Cordero, Harrison et al. (2021) and Amon et al. (2021) show the crucial validation of that choice.

Another important source of uncertainty is the shape calibration process, and particularly its response to blending MacCrann et al. (2020). The common way of describing such effects (Heymans et al. 2006, Huterer et al. 2006) is a simple rescaling of the two-point function model prediction

$$\xi_{\pm}^{ij} \rightarrow (1 + m^i)(1 + m^j) \times \xi_{\pm}^{ij}, \quad (28)$$

where again the indices  $i, j$  indicate redshift bins, and  $m$  is assumed to be redshift and scale independent within each bin. This is our fiducial parameterization. As explained in Sec. II C, this approach is an approximation to the more complete methodology set out in MacCrann et al. (2020) (see in particular their Sec. 6.5), which incorporates the redshift-dependent impact of blending on the source redshift distributions. Nevertheless, an extensive series of tests indicate that Eq. (28) is sufficient for the statistical power of DES Y3 (see Amon et al. 2021, MacCrann et al. 2020).

## F. Incorporating Shear Ratios

In the DES Y3 analysis, we also incorporate small-scale shear ratios (SR) at the likelihood evaluation level. *This likelihood is included in all our present constraints unless explicitly noted otherwise.* This concept and its DES application is outlined in more detail by Sánchez, Prat et al. (2021). Essentially, SR is an additional lensing-based data vector with 9 data points made up of measured galaxy-galaxy lensing ratios. It is nearly independent of cosmology and galaxy bias, but responds to (and thus constrains) redshift distributions and intrinsic alignments. Note that this particular analysis feature is different from Y1 and most current cosmic shear results, a detail that should be kept in mind in any comparison of our findings with previous results.

We use SR on small scales that are not used in the  $3 \times 2$ pt likelihood ( $< 6$  Mpc/ $h$ ), where uncertainties are dominated by galaxy shape noise, such that the likelihood can be treated as independent of that from the  $\xi_{\pm}$  (and indeed the full  $3 \times 2$ pt) data. On these scales the SR measurement is only very weakly dependent on the matter power spectrum, and so immune to uncertainties in modeling  $P_{\delta}$  on small scales (Sánchez, Prat et al. 2021). In this work, we only employ SR as obtained from the fiducial MAGLIM lens sample Porredon et al. (2021b), and marginalize over its required nuisance parameters. These are a free (linear) galaxy bias coefficient per lens bin (which are unconstrained by cosmic shear and have no impact on the final posteriors), as well as lens redshift parameters (three shift  $\Delta z_l$  and three width parameters  $\delta z_l$ ; see Porredon et al. 2021a), on which we have relatively tight priors. The main impact of including SR is an improvement in constraints on IAs and redshift nuisance parameters. Through internal degeneracies, this translates into significantly improved constraints on  $S_8$ , of around 30% in our fiducial analysis setting. Further discussion can be found in Sánchez, Prat et al. (2021) (Sec. 6 and Fig. 10) and Amon et al. (2021) (Sec. X and Fig. 11). In particular, those papers also assess the impact of utilizing REDMAGIC, an alternative DES Y3 lens sample (Cawthon et al. 2020, Pandey et al. 2021, Rozo et al. 2016), in order to obtain the shear ratios.

## G. Angular Scale Cuts

### 1. Baryons

The impact of baryons on the matter power spectrum on cosmological scales is a source of considerable uncertainty (van Daalen et al. 2014, Harnois-Déraps et al. 2015, Semboloni et al. 2013). While feedback processes from active galactic nuclei (AGN) and supernovae heat up the halo environment and tend to suppress matter clustering, metal enrichment can offer cooling channels that in fact increase power on small scales. These effects also depend on redshift and galaxy evolution (Semboloni et al. 2011).

As a fiducial strategy for mitigating the uncertainties coming from baryonic physics, we employ a gravity-only matter power spectrum (CAMB + HALOFIT) as described in Sec. IV C, and remove angular scales from the data vector which are signif-

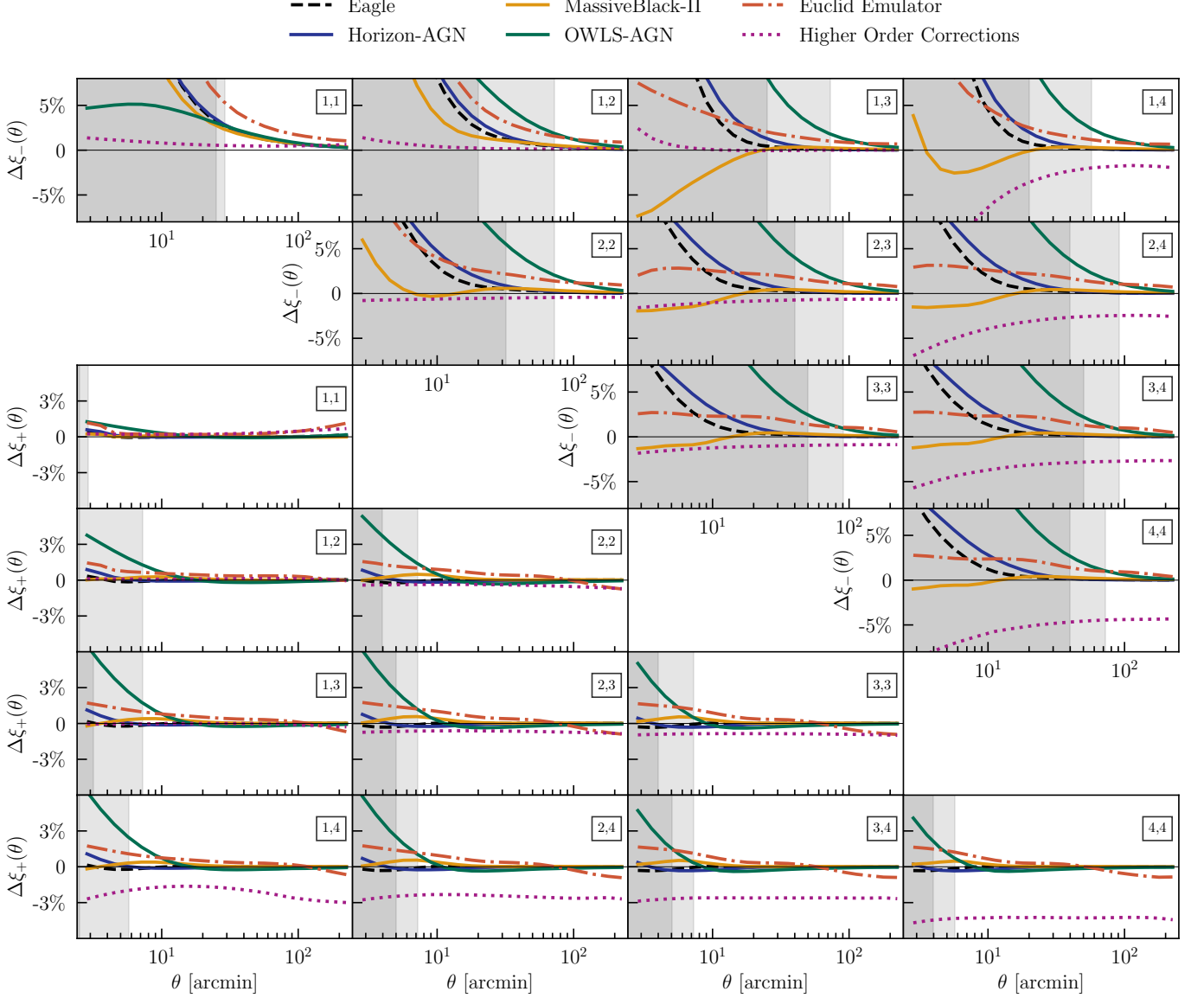


FIG. 5. The theoretical contribution of different modeling systematics to  $\xi_{\pm}$ ; for each systematic we show the fractional impact, relative to the fiducial case  $\Delta\xi_{\pm} \equiv (\xi_{\pm} - \xi_{\pm}^{\text{sys}})/\xi_{\pm}$ . Fiducial scale cuts are shown as light shaded bands and are derived jointly for cosmic shear,  $2 \times 2$ pt and  $3 \times 2$ pt. Darker bands correspond to (less stringent) scale cuts that are optimized for cosmic shear and  $3 \times 2$ pt in  $\Lambda$ CDM only. Error bars on  $\xi_{+(-)}$  are, at their smallest, around 15% (20%) of the signal, so the maximum contamination shown here  $\mathcal{O}(1 \sim 5\%)$  is still significantly below the sensitivity of our data. We also find that none of these forms of contamination project translate into a bias in cosmological parameters at Y3 precision, despite appearing coherent in some redshift bins. EAGLE (black dashed), HORIZON-AGN (solid blue), MASSIVEBLACK-II (solid yellow) and OWLS-AGN (solid green) represent scenarios for baryonic physics, and obtained from the power spectra of hydrodynamic simulations, while Euclid Emulator (dot-dashed red) modifies the non-linear gravity-only power spectrum. Higher Order Corrections (dashed purple) is the theoretical impact of reduced shear and source magnification.

icantly affected by feedback processes. We have additionally verified with synthetic data that marginalizing over baryonic halo model parameters with conservative priors does not lead to an actual gain in the  $S_8 \times \Omega_m$  sub-space. Specifically, we use HMCODE (Mead et al. 2015) and free both the halo concentration amplitude  $A$  with a flat prior [1.0, 7.5] and the bloating parameter  $\eta_0$  with a flat prior [0.4, 1.0]. These priors were chosen such that all baryonic scenarios tested in Mead et al. (2015) would be encapsulated by the allowed halo model modification. We find, with these fairly conservative priors, that attempting to model progressively smaller scales in synthetic data vectors results in tighter constraints on the baryon nuisance parameters  $A$  and  $\eta_0$ , while leaving the constraining power on  $S_8 \times \Omega_m$  approximately unchanged.

## 2. Determining scale cuts

We define scale cuts based on the  $\Delta\chi^2$  between noiseless synthetic cosmic shear data vectors, generated with and without a baryonic ‘‘contamination’’, according to Eq. (5). For the baryonic data vector, we use the OWLS (van Daalen et al. 2011, Schaye et al. 2010) matter power spectrum. The AGN-feedback implementation in this suite of simulations represents one of the most extreme scenarios in the literature, and thus characterizes the most conservative case for the baryonic contamination we expect our analysis to be safe against.

To determine scale cuts for the cosmic shear two-point functions  $\xi_{\pm}^{ij}$ , we evenly distribute the  $\Delta\chi^2$  threshold among tomographic bins, and limit each of the shear two-point angular functions:

$$\left( \xi_{\pm, \text{baryon}}^{ij} - \xi_{\pm, \text{base}}^{ij} \right)^t \mathbf{C}_{\text{sub}}^{-1} \left( \xi_{\pm, \text{baryon}}^{ij} - \xi_{\pm, \text{base}}^{ij} \right) < \frac{\Delta\chi_{\text{threshold}}^2}{N} \quad (29)$$

where  $N = 20$  is the number of cross-correlations between four tomographic redshift bins in the two shear correlations  $\xi_{\pm}$ . The  $\mathbf{C}_{\text{sub}}^{-1}$  sub-matrix of the inverse covariance matrix corresponds to the specific tomographic bin. For each element of  $\xi_{\pm}^{ij}$ , we find the minimum angular separation that satisfies Eq. (29) and exclude data points at smaller separations.

For a given threshold, we run two chains using the fiducial pipeline: one on a baseline (systematics-free) synthetic data vector and one on a contaminated synthetic data vector which includes the modification of the OWLS-AGN matter power spectrum. Our most important criterion for selecting a threshold is that the peaks of the baseline and contaminated posteriors in the 2-D parameter subspace of  $S_8 \times \Omega_m$  must be separated by less than  $0.3\sigma$ . We make this analysis additionally conservative by fixing redshift and shear uncertainties and the neutrino mass. It is also worth noting that scale cuts for the cosmic shear correlations  $\xi_{\pm}^{ij}$  are determined in tandem with the  $2 \times 2$ pt and  $3 \times 2$ pt analyses. That means a candidate cut passing the  $< 0.3\sigma$  threshold for cosmic shear is then combined with cuts from the analogous  $2 \times 2$ pt analysis, and the finally selected scale cuts for cosmic shear and  $2 \times 2$ pt are those which also satisfy  $3 \times 2$ pt  $< 0.3\sigma$ . In practice, this procedure ensures that the full DES Y3 analysis is optimized,

as opposed to each probe alone being optimally constraining (although see the discussion in Sec. IV G 3).

After an iterative procedure, we finally select the scale cuts that lead to a threshold of  $\Delta\chi^2 < 0.5$  for cosmic shear systematics. These cuts are shown in grey bands in Figures 2 and 5, and roughly correspond to bounds that are, at the smallest,  $\theta_{\text{min},+} = 2.4$  and  $\theta_{\text{min},-} = 30$  arcmin in  $\xi_+$  and  $\xi_-$  respectively, but differ significantly between redshift bin pairs. The large-scale limit of the data vector is 250 arcmin, chosen to match the DES Y1 analysis (Troxel et al. 2018). After cuts, we are left with 166 and 61 angular bins in  $\xi_+$  and  $\xi_-$ , or a total of 227 data points.

Fig. 5 demonstrates that, by making our data immune to the feedback impact of OWLS-AGN, we are also safely excluding parts of the data vector that are affected by systematic uncertainty due to the matter power spectrum modeling (Euclid Emulator; Euclid Collaboration 2019), baryonic feedback (MASSIVEBLACK-II, HORIZON-AGN and EAGLE; respectively Dubois et al. 2014, Khandai et al. 2015, Schaye et al. 2014) and higher order shear corrections including reduced shear and source magnification (Sec. V A 2 and Krause et al. 2021). See Sec. V for more details. It is important to note that error bars on  $\xi_+$  and  $\xi_-$  are, at their smallest, around 15% and 20% of the signal, so the deviations seen in Fig. 5 are well below the sensitivity of our data. We refer the reader to Krause et al. (2021) for an exploration of different baryonic power spectra, and in particular other extreme feedback cases such as ILLUSTRIS (Vogelsberger et al. 2014).

## 3. $\Lambda$ CDM Optimized scale cuts

As discussed above, our fiducial choice of scale cuts, by construction, optimizes the joint  $3 \times 2$ pt analysis of Dark Energy Survey Collaboration (2021a), which includes not just cosmic shear, but also galaxy clustering and galaxy-lensing. Those scale cut tests were also required to pass the same tolerance threshold in both  $\Lambda$ CDM and  $w$ CDM, making the exercise even more stringent. The end result is that the fiducial scale cuts used for Y3 cosmic shear (both in this work and in Amon et al. 2021) are more conservative than strictly necessary for the baseline cosmic shear-only  $\Lambda$ CDM analysis. For this reason we repeat a subset of our analyses with an alternative set of relaxed scale cuts, designed to maximize the constraining power in cosmic shear and  $3 \times 2$ pt  $\Lambda$ CDM alone (referred to as the ‘‘ $\Lambda$ CDM Optimized’’ analysis, disregarding  $2 \times 2$ pt-alone and  $w$ CDM). These optimized cuts increase the total number of data points from 227 to 273 (184  $\xi_+$  and 89  $\xi_-$ ). We discuss the improvement in constraining power that comes from this choice in Sec. VI. While the gain in angular scales is not uniform across redshift bins, the minimum scale is reduced by a factor of 20–70% for different bin pairs. We show the  $\Lambda$ CDM optimized scale cuts as darker grey bands on Fig. 5 (compared to the lighter grey bands which correspond to the fiducial scale cuts and eliminate more data points on small scales).

## H. Choice of Priors

The priors on cosmological and nuisance parameters are summarized in Table I. Our choice of cosmological priors is relatively conservative, in order to ensure that the posterior distributions can span any point in parameter space we believe to be reasonable. Although the argument can be made for imposing more informative priors on our less well constrained parameters (primarily  $\Omega_b$  and  $h$ ) based on external data (e.g. Planck), we choose not to do so here. The main reasoning here is that we want to maintain the statistical independence of the DES Y3 cosmic shear data, to avoid double-counting information when combining our results with external data sets.

In contrast, we *do* impose informative priors on the measurement systematics (i.e. shear bias and photo- $z$  error), which are derived from image simulations (MacCrann et al. 2020; Gatti, Sheldon et al. 2021) and tests using Buzzard (Myles, Alarcon et al. 2020; Buchs, Davis et al. 2019). More details on how these priors were chosen, and the dominant uncertainties can be found in those papers. It is worth bearing in mind that, while cosmic shear has only limited potential to self-calibrate photo- $z$  errors (Samuroff et al. 2017), the combination with galaxy clustering and galaxy-galaxy lensing breaks parameter degeneracies; in the current study we are prior dominated in these systematics parameters, but in the  $3 \times 2$ pt case this is significantly less true (Dark Energy Survey Collaboration 2021a).

We choose uninformative priors on all of the five IA parameters, motivated by the relative lack of constraints reported in the literature for the TATT model (with the exception of Samuroff et al. 2019). Since IAs are highly sensitive to the composition of the galaxy population, as well as physical characteristics like redshift and luminosity, deriving appropriate priors is a complicated matter. We consider whether these IA parameters are detected by the data in Sec. VIC.

It is finally worth remarking that marginalizing posterior distributions from a high-dimensional parameter space down to 1D is prone to projection effects that can significantly displace the full-dimensional best-fit from the mean (or peak) of the 1D parameter distributions. The magnitude of these effects can be non-trivially affected by the choice of priors, particularly on the parameters that are not well constrained by cosmic shear data alone. We again refer the reader to Krause et al. (2021), where tests of these effects using the priors given in Table I can be found.

## V. MODEL ROBUSTNESS TESTS ON SYNTHETIC DATA

We now test the assumptions of the model set out in the previous section using contaminated mock data. The idea is that we generate data with astrophysical effects that are *not contained* in the model, and seek to quantify the level of bias they can cause to parameters of interest. These tests fall into two different categories, both involving synthetically generated  $\xi_{\pm}$ : noiseless analytic data (which is produced using the theory pipeline), and simulated data (measured using mock catalogs

Parameter	Prior
<b>Cosmological Parameters</b>	
$\Omega_m$	U[0.1, 0.9]
$A_s$	U[0.5, 5.0] $\times 10^{-9}$
$\Omega_b$	U[0.03, 0.07]
$n_s$	U[0.87, 1.07]
$h$	U[0.55, 0.91]
$\Omega_b h^2$	U[0.6, 6.44] $\times 10^{-3}$
$w$	U[-2, -0.333]
<b>Calibration Parameters</b>	
$m_1$	$\mathcal{N}(-0.0063, 0.0091)$
$m_2$	$\mathcal{N}(-0.0198, 0.0078)$
$m_3$	$\mathcal{N}(-0.0241, 0.0076)$
$m_4$	$\mathcal{N}(-0.0369, 0.0076)$
$\Delta z_1$	$\mathcal{N}(0.0, 0.018)$
$\Delta z_2$	$\mathcal{N}(0.0, 0.015)$
$\Delta z_3$	$\mathcal{N}(0.0, 0.011)$
$\Delta z_4$	$\mathcal{N}(0.0, 0.017)$
<b>Intrinsic Alignment Parameters</b>	
$a_1$	U[-5, 5]
$a_2$	U[-5, 5]
$\eta_1$	U[-5, 5]
$\eta_2$	U[-5, 5]
$b_{TA}$	U[0, 2]
<b>Shear Ratio Parameters</b>	
$\Delta z_1^{\text{lens}}$	$\mathcal{N}(-0.009, 0.007)$
$\Delta z_2^{\text{lens}}$	$\mathcal{N}(-0.035, 0.011)$
$\Delta z_3^{\text{lens}}$	$\mathcal{N}(-0.005, 0.006)$
$\delta_{z,1}^{\text{lens}}$	$\mathcal{N}(0.975, 0.062)$
$\delta_{z,2}^{\text{lens}}$	$\mathcal{N}(1.306, 0.093)$
$\delta_{z,3}^{\text{lens}}$	$\mathcal{N}(0.870, 0.054)$
$b_g^{1-3}$	U[0.8, 3]

TABLE I: A summary of the priors used in the fiducial Y3 analysis. The top seven rows are cosmological parameters, while those in the lower sections are nuisance parameters corresponding to astrophysics and data calibration. We fix  $w = -1$  when analyzing  $\Lambda$ CDM. Priors are either uniform (U) or normally-distributed,  $\mathcal{N}(\mu, \sigma)$ .



from N-body simulations). Chronologically, all of the tests described in this Section were performed *before unblinding*.

### A. Analytically generated data

We analyze these data with the fiducial model, priors and analysis choices (including scale cuts), to verify that the inferred cosmological constraints are biased by less than  $0.3\sigma$  in the  $S_8 \times \Omega_m$  plane. Fig. 5 shows the systematic contributions to the contaminated data vectors for a number of the tests described below.

#### 1. Nonlinear matter power spectrum

By construction, the scale cuts defined in Sec. IV G are conservative and mitigate most of the impact of small-scale modeling uncertainties in the 3-dimensional gravity-only matter power spectrum. To show that these uncertainties are negligible, we generate a synthetic cosmic shear data vector using the gravity-only Euclid Emulator (Euclid Collaboration 2019). Though it can be used at a relatively restricted range of cosmologies, the emulator is the most accurate method for estimating the nonlinear  $P_\delta(k)$  currently available to the community. The deviation between this data vector and the fiducial prediction is shown in Fig. 5 (dot-dashed red), and is at most  $\sim 1$  percent on the scales included in our analysis. We verify that the bias in cosmological parameters when analysing these mock data with the fiducial  $\Lambda$ CDM model is below  $0.3\sigma$  (Krause et al. 2021).

In addition to the test of the nonlinear growth model, we also generate synthetic data vectors using matter power spectra derived from the HORIZON-AGN, EAGLE and MASSIVEBLACK-II hydrodynamic simulations (Fig. 5, Dubois et al. 2016, Khandai et al. 2015, Schaye et al. 2015). The  $P_\delta$  obtained from these simulations contain the imprint of baryonic feedback processes. Although these effects are notoriously difficult to model, the simulations are useful as a measure of the range of uncertainty in the baryonic contributions. That is, if we can demonstrate insensitivity to a range of (at least semi-) realistic scenarios, then that offers some reassurance that our cosmological results are unaffected. We confirm that the resulting biases are significantly smaller than  $0.3\sigma$  (Krause et al. 2021). At some level this result is expected, since our fiducial scale cuts are constructed based on the OWLS-AGN model, which is a relatively extreme feedback scenario. It does, however, demonstrate that our scale cut prescription is generally conservative, and is not fine-tuned to the specific redshift or scale dependence of OWLS-AGN.

#### 2. Higher-order shear contributions

We also verify that the effects of reduced shear, source sample magnification and source clustering do not affect the cosmic shear constraints significantly. These generally enter as higher-order corrections. For the case of reduced shear

and source magnification, the correction to the convergence angular power spectrum has the form (Dodelson et al. 2006, Schmidt et al. 2009, Shapiro 2009):

$$\Delta C_k^{ij}(\ell) = 2(1 + C_s^i) \int \frac{d^2\ell_1}{(2\pi)^2} \cos 2\phi_{\ell_1} B_k^{ij}(\vec{\ell}_1, \vec{\ell} - \vec{\ell}_1, -\vec{\ell}) \quad (30)$$

where  $B_k^{ij}$  is the convergence bispectrum,  $C_s^i$  is a magnification coefficient related to the slopes of the flux and size distributions (Schmidt et al. 2009) and  $\phi_{\ell_1}$  is the angle between modes  $\vec{\ell}$  and  $\vec{\ell}_1$ . We show in Fig. 5 this contribution to the shear power spectrum for the source magnification coefficients  $C_s^{1\dots 4} = (-1.17, -0.64, -0.55, 0.80)$  as measured by Elvin-Poole et al. (2021) in each redshift bin. We note that the theoretical modeling of non-linear scales in the matter bispectrum is a dominant source of uncertainty in the calculated correction (see for instance Lazanu et al. (2016)). In particular, we utilize the non-linear (HALOFIT) power spectrum in the tree-level bispectrum description, but future analyses will possibly benefit from improved approaches, such as B-HALOFIT (Takahashi et al. 2020). Still, the largest potential biases as seen in Fig. 5 are  $< 5\%$ , smaller than our error bars ( $\gtrsim 15\%$ ). We also find, like Prat et al. (2021) demonstrate for galaxy-galaxy lensing, that the reduced shear contribution to cosmic shear as computed perturbatively overestimates the impact of that correction when compared to measurements of reduced shear on BUZZARD simulations. The leading sources of higher-order corrections and their contribution for DES Y3 are computed in more detail in Krause et al. (2021). Their results confirm that these effects are negligible in terms of parameter constraints, and provide a guide for future DES and next generation analyses.

#### 3. Intrinsic Alignments

Prior to unblinding of the real data, we carried out a series of tests designed to verify the robustness and sufficiency of our fiducial choice of intrinsic alignment model. Our expectation is to find, using synthetic data, a set of model choices that prevent biases coming from an incorrect characterization of IA. To this end, we generate two simulated data vectors: one including full TATT contributions with input parameters given by the mean posterior IA constraints from the Y1  $3 \times 2$ pt analysis of Samuroff et al. 2019 ( $a_1 = 0.7, a_2 = -1.36, \eta_1 = -1.7, \eta_2 = -2.5, b_{TA} = 1$ ), and another with the NLA subset of those parameters ( $a_1 = 0.7, \eta_1 = -1.7, a_2 = 0, b_{TA} = 0$ ). Our results can be summarized as:

- Fitting the NLA data vector with either the NLA or TATT model results in no significant biases. Projection effects in the 1D  $S_8$  posterior are negligible in both cases. The more complex model weakens the constraint by increasing the  $S_8$  error bar by 14%. See Appendix C for a discussion of this on real data, where we find a somewhat similar loss in constraining power.

- Likewise, and as expected, fitting the TATT data vector using TATT recovers the input cosmology and IA parameters.
- Fitting the TATT data vector with the simpler NLA model results in significant residual bias in  $\sigma_8$ ,  $S_8$  and  $\Omega_m$ . We also see artificially tightened constraints in this case, as the posteriors on some parameters begin to hit the prior edges, and a significantly degraded best-fit  $\chi^2$ .

The findings above are summarized in Fig.6, and a more detailed discussion of these tests, and what they reveal about the systematic error budget, can be found in Appendices B and C. The configuration which presented significant biases (NLA modeling of TATT synthetic data) should be considered as a sufficiency test of the simpler IA model, which NLA *fails to meet at the synthetic data level*. This means that, given our expectation from the DES Y1 results of Samuroff et al. (2019), and given our deliberate choice of relying solely on synthetically generated data to make a model selection at the blinded analysis stage, our findings provide a strong argument for adopting TATT as our fiducial choice. In summary, previous results on DES Y1 data do not allow us to rule out alignments at the level present in the simulated TATT data vector used in these tests, and we thus choose TATT as our fiducial model. We explore in Sec. VIC 2 whether this pre-unblinding expectation of the inadequacy of NLA to fit our data is actually realized and what level of IA complexity our data seems to point to. Foreshadowing that discussion, post-unblinding we find NLA to be compatible with DES Y3 data, and that our results for  $S_8$  are robust to the choice of the IA model.

Our fiducial setup makes the explicit assumption that the redshift evolution of IAs in DES Y3 can be adequately described by a power law (Eq. 24 and 25). To test this particular analysis choice, we produce synthetic data using the recipe set out in (Krause et al. 2016; see their Sec. 4.1). In brief, we assume a particular luminosity scaling for the linear alignment strength  $\beta_1$ , which is combined with a measured luminosity function from the Deep Extragalactic Evolutionary Probe2 (DEEP2; Faber et al. 2007) spectroscopic survey, and the measured red fraction  $f_{\text{red}}(z)$ , to predict an IA signal with a non-trivial redshift dependence. Both the redshift and luminosity evolution of red and red+blue galaxies are thus extrapolated (not measured) down to the limiting magnitude over the redshift range of the DES sample. Our baseline is to use the best fitting  $\beta_1$  from Joachimi et al. (2011) and assume zero alignments in blue galaxies. Note that while constraints on the luminosity dependence of the  $a_2$  component of the TATT model do exist in the literature, they are still relatively weak, and consistent with  $\beta_2 = 0$ ; for this reason, we restrict ourselves to the NLA case. The results are described in more detail in Krause et al. (2021), which also includes similar tests in the context of  $2 \times 2$ pt and  $3 \times 2$ pt analyses. With our fiducial input parameters, we find biases of up to  $0.5\sigma$ . Although larger than our  $0.3\sigma$  passing requirement, it is worth bearing in mind that this test relies on a number of assumptions and extrapolations and is quite possibly unrealistically stringent. Given this, and in the absence of a stronger test, we maintain

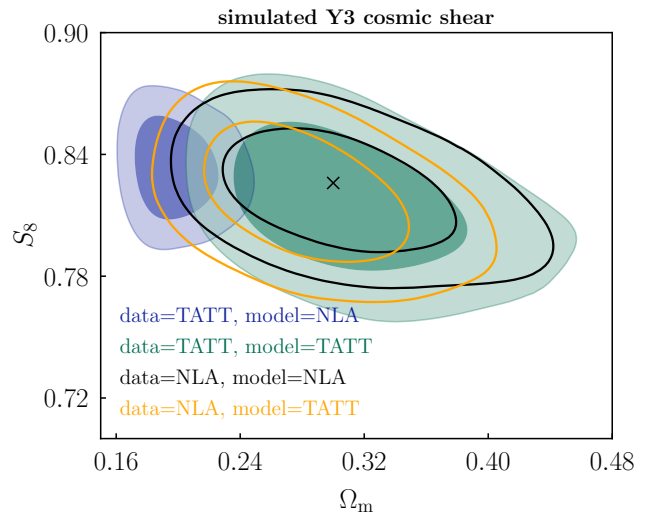


FIG. 6. Posteriors from the analysis of *synthetic data vectors with differing input IA signals*, with a known input cosmology (black cross). Strikingly, if IAs in DES Y3 are present at the level of those in the synthetic TATT data vector (chosen to match the DES Y1 result of Samuroff et al. (2019)), the NLA model is not sufficient to recover the true cosmology (blue contours). We explore whether this scenario is actually reproduced in real data in Sec. VIC 2. All other model/data combinations recover the input cosmology to within significantly less than  $1\sigma$ .

our fiducial model and decide to test for the redshift scaling of the  $a_1$  parameter *a posteriori*, after unblinding, in Sec. VII.

Finally, we assess whether the interaction between erroneous calibration of photo- $z$ s and their uncertainties and our IA model can lead to biased estimates of cosmology. A rationale for this test is that photo- $z$  mischaracterization can produce amplitude shifts in cosmic shear correlations. These shifts can potentially be degenerate with IA model parameters, and bias their posterior distributions, which in turn can lead to biases in cosmology as discussed, e.g., by Wright et al. (2020). Another example is when  $n(z)$  distribution tails are underestimated in the data, leading to an erroneous assumption of small overlap between redshift bins and consequently a suppressed II contribution, which is more sensitive to this overlap than the lensing signal is (see also Appendix C of Fortuna et al. 2021).

We seek to test our robustness to such redshift-IA interactions by reanalysing the synthetic TATT data vector described above repeatedly using alternative (incorrect) redshift distributions in the modeling. For this test, we use the fiducial analysis setup including redshift error as parameterised by free  $\Delta z$  shift parameters. Rather than introducing analytic distortions to the fiducial  $n(z)$ , we draw alternative realizations from the ensemble generated by the SOMPO pipeline (Myles, Alarcon et al. 2020). This should naturally capture plausible cases of redshift error, beyond what can be captured by a simple shift in the mean  $z$ . We select three redshift realizations representative of a range of cases in terms of the change in the predicted data vector ( $\Delta\chi^2$ ). The results are shown in more detail in Appendix B, but in brief we find only small ( $< 0.3\sigma$ ) shifts in  $S_8$  in the different scenarios. Even in the most extreme case,

where  $\Delta\chi^2 \sim 14$ , the input IA parameters and cosmology are accurately constrained. In summary, at the level of synthetically generated data with plausible fluctuations in  $n(z)$ , we do not see evidence that the relatively flexible IA model is absorbing redshift error to a degree that could bias cosmology or IA parameters.

## B. Mock data from N-body simulations

While most model tests are performed using noiseless analytic data vectors, a subset of tests make use of mock catalogs from N-body simulations. We employ two sets of simulations: BUZZARD (DeRose et al. 2019a) and the MICE-Grand Challenge Galaxy and Halo Light-cone catalog<sup>13</sup> (Fosalba et al. 2015b). With BUZZARD, a thorough testing of the DES Y3 measurement, photo- $z$  and likelihood pipelines is performed not only for cosmic shear but also for the other DES  $3 \times 2$ pt probes (DeRose et al. 2021a). With MICE, we focus specifically on testing the IA model.

### 1. Buzzard v2.0

The BUZZARD v2.0 simulations (BUZZARD hereafter) are a suite of 18 simulated galaxy catalogs built on  $N$ -body lightcone simulations (DeRose et al. 2019a; DeRose et al. 2021b; Wechsler et al. 2021). The  $N$ -body simulations are produced using the L-GADGET2  $N$ -body code, a memory optimized version of GADGET2 (Springel 2005). The initial conditions for the simulations are generated at  $z = 50$  using 2LPTIC and linear power spectra computed by CAMB at the BUZZARD flat  $\Lambda$ CDM cosmology:  $(\sigma_8, n_s, h, \Omega_m, \Omega_b) = (0.82, 0.96, 0.7, 0.286, 0.046)$ . Each simulation is run on three different and independent boxes with sizes  $(1.0, 2.6 \text{ and } 4.0)^3 h^{-3} \text{Gpc}^3$  containing  $(1400, 2048 \text{ and } 2048)^3$  particles respectively, and a lightcone with footprint area over  $10,000 \text{ deg}^2$  is produced from each of the sets. Galaxies are included in these lightcones with properties such as position, ellipticity and spectral energy distribution using the ADDGALS algorithm, with details specified in Wechsler et al. (2021).

Weak lensing quantities are introduced via ray-tracing (Becker 2013), and the complete  $3 \times 2$ pt cosmology analysis of DES Y3 is reproduced in these simulations (DeRose et al. 2021a). For the validation of cosmic shear, BUZZARD provides a way to verify that several astrophysical effects ignored in our modeling are, as expected, sub-dominant. In addition to higher order corrections from reduced shear and source sample magnification which are also examined in this paper, BUZZARD mocks contain source galaxy clustering, anisotropic redshift distributions across the survey area and multiple-plane lensing deflections. We point the reader to DeRose et al. (2021a) for more details.

### 2. Intrinsic Alignments Testing with MICE

The MICE Grand Challenge run is an  $N$ -body gravity-only simulation with  $4096^3$  collisionless particles with masses  $2.927h^{-1}M_\odot$  in a  $3072h^{-1} \text{Mpc}$  box using a flat  $\Lambda$ CDM cosmology with  $(\sigma_8, n_s, h, \Omega_m, \Omega_b) = (0.8, 0.95, 0.7, 0.25, 0.044)$ . Dark matter halos are identified as friends-of-friends groups and populated with synthetic galaxies using a hybrid method of Halo Occupation Distribution modeling and Abundance Matching (Carretero et al. 2015, Crocce et al. 2015, Fosalba et al. 2015a,b, Hoffmann et al. 2015).

Although MICE does not account for baryonic feedback and other hydrodynamic processes related to galaxy formation and evolution, an intrinsic galaxy alignment signal is implemented in the mock using a semi-analytic technique (Appendix A; see Hoffmann et al. 2021 for a detailed overview). In brief, shapes and orientations are assigned to galaxies based on a combination of their color and luminosity, and the spin and orientation of the host halos. The model is conceptually similar to those of Okumura et al. (2009) and Joachimi et al. (2013a,b), and has been tuned to match the distribution of galaxy axis ratios from COSMOS observations (Laigle et al. 2016, Scarlata et al. 2007) in bins of redshift, stellar mass and color. Various diagnostic tests have been carried out, and it has been shown to reproduce the projected intrinsic galaxy-shape correlation  $w_{g+}$  of luminous red galaxies in the spectroscopic BOSS LOWZ sample (Singh & Mandelbaum 2016) as a function of luminosity (Hoffmann et al. 2021).

The MICE IA mock is particularly valuable for our purposes, since it provides an alternative IA prediction with its own redshift dependence, which does not assume either NLA or TATT. At the time of writing, however, only a single realization of the IA mock is available, and we thus we cannot apply the more stringent criteria of DeRose et al. (2021a) for validating the DES Y3 pipeline on simulations. For this reason, passing the tests presented in this subsection was not an unblinding prerequisite.

We describe in Appendix A how a source sample is created with the MICE mock which reproduces basic DES Y3 specifications. The measurement of cosmic shear correlation functions  $\xi_\pm$  is made using TREECORR in a similar way to on the real Y3 data. Since the MICE catalog contains the gravitational shears  $\gamma_{1,2}$  (G) and the intrinsic shapes  $\epsilon_{1,2}$  (I), we can estimate the GG, GI and II components separately if desired. We thus measure two data vectors: a version which contains the GG signal only (the ‘‘baseline’’) and another version in which G+I shears are summed together (Seitz & Schneider 1997 Eq. 3.2). In this second case, the two-point correlations correspond to GG+GI+II (a ‘‘contaminated’’ data vector including the noisy IA signal). With the shape noise and effective number density of MICE galaxies per redshift bin determined in Appendix A, we construct an analytic covariance using CosmoLIKE (Krause & Eifler 2017), but rescale the shape noise terms so that the figure-of-merit of the fiducial model constraints in the  $S_8 \times \Omega_m$  plane is similar to that of the other simulated cosmic shear tests. We employ the fiducial scale cuts and nuisance parameters.

<sup>13</sup> <https://cosmohub.pic.es/home>

With the same likelihood pipeline used in other tests, we find that the bias (distance between posterior peaks) in the  $S_8 \times \Omega_m$  plane between the baseline and contaminated data vectors is  $0.6\sigma$  for the fiducial (TATT) IA model. A similar run, but using NLA as the model, results in a  $0.3\sigma$  bias between baseline and contaminated data. Since the more flexible TATT model should, in principle, be able to capture any IA scenario that NLA can, we ascribe this difference to competing shifts due to the statistical fluctuations and the simpler IA model, which partially cancel. Furthermore, the posteriors from  $a_1$  and  $\eta_1$  coming from the two models are fully consistent.

One important caveat here is that we have only a single noise realization; the size of the biases quoted for NLA and TATT therefore cannot be decoupled from fluctuations due to shape noise and cosmic variance. This point complicates the interpretation of our results. That said, there is a qualitative difference between the MICE+IA simulated data vectors and the other synthetic  $\xi_{\pm}$  vectors: while all other data are analytically generated using either the NLA or TATT model, the MICE data are agnostic with respect to the analytic model and are tuned to fit observed IA signals (albeit at significantly lower redshifts). It is, then, reassuring that we obtain reasonable IA constraints, and the cosmological parameters  $S_8$  and  $\Omega_m$  are not catastrophically biased. Analysis of a greater number of MICE+IA mocks, in order to disentangle IAs from noise and cosmic variance, is an interesting extension left for future work, but beyond the scope of this paper.

## VI. PARAMETER CONSTRAINTS

In this section we present the baseline cosmic shear constraints from DES Y3. Sec. VIA presents our main results in  $\Lambda$ CDM, Sec. VIB shows we find no detection of the dark energy equation of state parameter  $w$ , and Sec. VIC2 presents a detailed comparison of the evidence for different IA models. A series of extended models, including parameterized deviations from General Relativity, and additional neutrino species, will be discussed in [Dark Energy Survey Collaboration 2021b](#), which is in preparation. A number of external data sets are introduced in this section, details of which can be found in Sec. VIIIA. For a summary of our fiducial results, and of the robustness tests discussed in the next section, see Table II and Fig. 10. To facilitate comparison of constraining power, we also show the 2D  $S_8 - \Omega_m$  Figure of Merit (FoM), which is defined as  $\text{FoM} \equiv \det\{C_{S_8, \Omega_m}\}^{-1/2}$ , where  $C_{S_8, \Omega_m}$  is a sub-block of the parameter covariance matrix.

### A. $\Lambda$ CDM

The posteriors from our fiducial analysis are shown in Fig. 7, along with our  $\Lambda$ CDM optimized result (with scale cuts that maximize the constraining power in  $\Lambda$ CDM, see Sec. IVG; black dashed), alongside Planck 2018 TT+TE+EE+lowE (no lensing, [Planck Collaboration 2020a](#); yellow). As with almost all previous lensing analyses to date, DES Y3 favors a somewhat lower  $S_8$ , a pattern we discuss further in Sec. VIIIB.2.

The marginalized mean  $S_8$  and  $\Omega_m$  values in  $\Lambda$ CDM are:

$$\begin{aligned} S_8 &= 0.759^{+0.023}_{-0.025} & (0.755) & (\Lambda\text{CDM}) \\ \Omega_m &= 0.290^{+0.039}_{-0.063} & (0.293) & (\Lambda\text{CDM}) \end{aligned}$$

where uncertainties are 68% confidence intervals. These 1D constraints represent our best estimate for the uncertainties within our analysis framework. It is worth bearing in mind, however, that comparison of constraining power between different lensing surveys is complicated by differences in the analysis choices and priors ([Chang et al. 2019](#)). Our fiducial  $\Lambda$ CDM analysis has 222 effective degrees of freedom<sup>14</sup>. At the maximum posterior (MAP) point we obtain  $\chi^2/\text{dof} = 237.7/222 = 1.07$ , and a corresponding  $p$ -value of 0.22.

One important question arising from Fig. 7 is the extent to which the DES and Planck results are consistent with each other. Given that we are considering a complex 28 dimensional parameter space, assessing agreement purely using projected contours can be misleading (see [Lemos et al. 2020](#) for discussion). We present a more rigorous quantitative discussion of possible tensions with external data sets in Sec. VIIIB.

Different angular scales of the cosmic shear correlation function have slightly different sensitivities to cosmological parameters ([Jain & Seljak 1997](#)). The degeneracy between  $\sigma_8$  and  $\Omega_m$  is such that the best constrained combination is  $S_8 = \sigma_8(\Omega_m/0.3)^\alpha$  with  $\alpha \approx 0.5$ , but we can also determine the exponent  $\alpha$  directly from the data. We carry out a Principal Component Analysis of the projected  $\Lambda$ CDM posteriors to obtain the exponent value that most effectively decorrelates  $\sigma_8$  and  $\Omega_m$ . We find in our fiducial  $\Lambda$ CDM TATT analysis that  $\alpha = 0.586$ . To avoid confusion, we call the corresponding lensing amplitude  $\Sigma_8$  and find:

$$\Sigma_8 \equiv \sigma_8(\Omega_m/0.3)^{0.586} = 0.756^{+0.021}_{-0.021} \quad (0.729) \quad (\Lambda\text{CDM}).$$

Also shown in Fig. 7 (black dashed) are the results of our optimized cosmic shear analysis. The details of, and justification for, this additional analysis can be found in Sec. IVG3, but the key idea is to use a set of scale cuts that are tuned to maximise the constraining power of cosmic shear alone in  $\Lambda$ CDM. The result is a tighter constraint in the  $S_8 - \Omega_m$  plane in  $\Lambda$ CDM:

$$S_8 = 0.772^{+0.018}_{-0.017} \quad (0.774) \quad (\Lambda\text{CDM Optimized}).$$

As can be seen in Fig. 7, the gain in constraining power is asymmetric about the posterior peak, which has the effect of shifting the mean  $S_8$  up slightly. We consider the impact of this in terms of statistical consistency in Sec. VIIIB. The extra data points increase the number of effective degrees of freedom to 268, giving a goodness-of-fit at the maximum posterior of  $\chi^2/\text{dof} = 285.0/268 = 1.06$ , with a  $p$ -value of 0.22, similar to the fiducial analysis.

<sup>14</sup> We calculate the effective number of d.o.f. following [Raveri & Hu \(2019a\)](#), as the number of data points for  $\xi_{\pm}$  after scale cuts (227) minus the number of effective parameters ( $N_{\text{p,eff}} \approx 5$ ), given by  $N_{\text{p,eff}} \equiv N_{\text{p,nom}} - \text{Tr}\{C_{\Pi}^{-1}C_p\}$ , where  $N_{\text{p,nom}}$  is the nominal number of free parameters and  $C_{\Pi}$ ,  $C_p$  are respectively the covariances of the prior and posterior. As  $N_{\text{p,eff}}$  can change with different modeling choices, we re-compute it whenever necessary.

	$S_8$	$\hat{S}_8$	$\sigma_8$	$\Omega_m$	FoM $_{S_8\Omega_m}$	$\chi^2/\text{dof}$
<b>Fiducial <math>\Lambda</math>CDM</b>	$0.759^{+0.023}_{-0.025}$	0.755	$0.783^{+0.073}_{-0.092}$	$0.290^{+0.039}_{-0.063}$	927	237.7/222 = 1.07
$\Lambda$ CDM-Optimized	$0.772^{+0.018}_{-0.017}$	0.774	$0.795^{+0.072}_{-0.076}$	$0.289^{+0.036}_{-0.056}$	1362	285.0/268 = 1.06
No IAs	$0.774^{+0.017}_{-0.018}$	0.760	$0.775^{+0.071}_{-0.077}$	$0.306^{+0.040}_{-0.061}$	1253	243.3/225 = 1.08
NLA	$0.773^{+0.020}_{-0.021}$	0.773	$0.791^{+0.070}_{-0.086}$	$0.293^{+0.039}_{-0.056}$	1163	242.1/224 = 1.08
NLA, free $a_1$ per $z$ -bin	$0.790^{+0.022}_{-0.020}$	0.783	$0.834^{+0.075}_{-0.082}$	$0.275^{+0.035}_{-0.051}$	1144	246.3/221 = 1.11
$a_1 > 0$ prior	$0.755^{+0.022}_{-0.022}$	0.727	$0.731^{+0.064}_{-0.085}$	$0.327^{+0.046}_{-0.066}$	881	238.7/222 = 1.07
Fixed neutrino mass	$0.772^{+0.023}_{-0.023}$	0.748	$0.816^{+0.071}_{-0.094}$	$0.275^{+0.040}_{-0.052}$	1063	238.3/222 = 1.07
$w$ CDM cosmology	$0.735^{+0.023}_{-0.041}$	0.699	$0.723^{+0.071}_{-0.100}$	$0.319^{+0.050}_{-0.071}$	497	237.5/222 = 1.07
HMCode power spectrum	$0.772^{+0.026}_{-0.027}$	0.791	$0.793^{+0.088}_{-0.102}$	$0.294^{+0.038}_{-0.070}$	827	236.8/222 = 1.06
NLA, $\Lambda$ CDM-Optimized, fixed neutrino mass	$0.788^{+0.017}_{-0.016}$	0.775	$0.825^{+0.078}_{-0.79}$	$0.279^{+0.036}_{-0.053}$	1501	288.1/270 = 1.07
DES Y1	$0.780^{+0.027}_{-0.021}$	-	$0.764^{+0.069}_{-0.072}$	$0.319^{+0.044}_{-0.062}$	625	227/211 = 1.08
KiDS-1000 COSEBIs	$0.759^{+0.024}_{-0.021}$	-	$0.838^{+0.140}_{-0.141}$	$0.246^{+0.101}_{-0.060}$	650	85.5/70.5 = 1.21
HSC Y1 $C_\ell$	$0.780^{+0.030}_{-0.033}$	-	-	$0.162^{+0.086}_{-0.044}$	461	45.4/53 = 0.86
HSC Y1 $\xi_\pm$	$0.804^{+0.032}_{-0.029}$	-	$0.766^{+0.110}_{-0.093}$	$0.346^{+0.052}_{-0.100}$	402	162.3/167 = 0.97
Planck 2018 TT + TE + EE + lowE	$0.827^{+0.019}_{-0.017}$	-	$0.793^{+0.024}_{-0.009}$	$0.327^{+0.008}_{-0.017}$	3938	-

TABLE II: A summary of cosmological constraints from DES Y3 cosmic shear. In each case in the top 9 rows, we show the posterior mean and 68% confidence bounds on each  $S_8$  and  $\Omega_m$ , as well as the Maximum Posterior  $S_8$  value (denoted  $\hat{S}_8$ ), the 2D  $S_8 - \Omega_m$  Figure of Merit and the goodness-of-fit. These top rows also include, as a default, the shear ratio likelihood. A visual summary of the  $S_8$  constraints can be seen in Fig. 10. Constraints from KiDS-1000 and HSC are as nominally reported and were not re-processed under the DES Y3 priors and analysis choices, while Planck 2018 numbers have been obtained under DES Y3 cosmology priors.

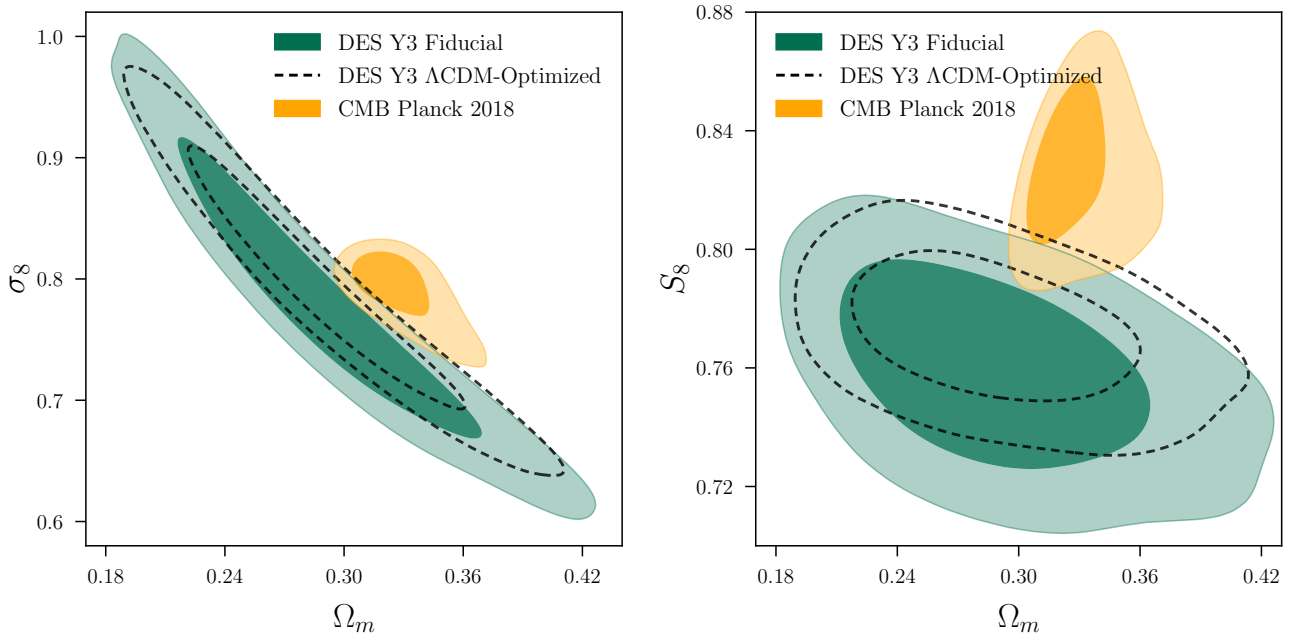


FIG. 7. The posteriors of DES Y3 cosmic shear (green) and Planck 2018 (TT+TE+EE+lowE no lensing; yellow). The shaded green contours are the results from our fiducial analysis, as described in Sec. IV, and constraints using scale cuts that are optimized for a  $\Lambda$ CDM-only analysis are shown in dashed lines. In each case we show both 68% and 95% confidence levels.

## B. $w$ CDM

A simple extension to  $\Lambda$ CDM is to free the dark energy equation-of-state parameter (previously fixed to  $w = -1$ ). The prior bounds on  $w$  in Table I are chosen such that cosmic acceleration is ensured by  $w < -1/3$ , and “phantom” models with  $-2 < w < -1$  are allowed. To assess the statistical preference of our data for this extended model, we evaluate the evidence ratio (see Sec. IV A) between  $\Lambda$ CDM and  $w$ CDM and find

$$R_{w/\Lambda} = 0.94 \pm 0.22,$$

which, based on the Jeffreys scale, is inconclusive in terms of preference for the model with free  $w$ . We thus find no evidence for (or against) a departure from  $\Lambda$ CDM using DES Y3 cosmic shear data alone.

While  $w$  is unconstrained by cosmic shear data within our priors, we can still constrain  $S_8$  in that parameter space, and we show its marginalized value in Fig. 10. We find a shift towards lower lensing amplitude, qualitatively implying that this model extension would not help to provide a solution to the differences with respect to Planck.

## C. Intrinsic Alignments

In addition to the main cosmological results, it is also interesting to consider what our data can tell us about IAs. Since we see no evidence that redshift error is biasing our IA results (see Appendix B), the IA parameters have a physical interpretation, and are potentially useful for future lensing analyses. For details about how changes in the IA model can affect the inferred cosmology see Sec. VII B.

### 1. IA Model Constraints

In Fig. 8 we show the IA posteriors in both TATT and NLA, as well as in the optimized  $\Lambda$ CDM analysis and NLA with a fixed redshift evolution parameter (for the numerical values, see Table III). Note that there are additional parameters in the full model (redshift indices and  $b_{\text{TA}}$ ), which are only weakly constrained and so are not shown here (see Appendix D for the full posterior). In all model scenarios, the data favor a smaller IA signal than DES Y1, although the values are consistent within the uncertainties (see Samuroff et al. 2019’s Fig. 12), resulting in a total GI+II contribution that is only at the level of a few percent of the cosmological signal (as shown in Fig. 2). In the fiducial TATT case, the data favor mildly negative  $a_1$ , combined with  $a_2 > 0$ :

$$\begin{aligned} a_1 &= -0.47^{+0.30}_{-0.52} \quad (-0.73) && \text{(TATT } \Lambda\text{CDM)} \\ a_2 &= 1.02^{+1.61}_{-0.55} \quad (1.88) && \text{(TATT } \Lambda\text{CDM)}. \end{aligned}$$

We observe asymmetric error bars in the posteriors of  $a_1$  and  $a_2$  and also find slight evidence for a secondary peak with reversed signs of the parameters. This hint of bimodality in the

IA subspace can be readily understood within our model: contributions to power spectra in equations (21)–(23) that scale as  $a_1^2$ ,  $a_2^2$  or  $a_1 a_2$  are insensitive to positive/negative sign flips. This degeneracy is further enabled in our present regime, where IA signals are small compared to the measurement errors, and terms proportional to  $a_1$  and  $a_2$  partially cancel each other when they have opposite signs. We also note that the secondary peak is less pronounced in the optimized analysis shown in Fig. 8, and Amon et al. (2021) shows that the addition of the shear ratios data in our fiducial constraints has already contributed to suppressing the bimodal feature. We thus interpret this doubly-peaked posterior as an internal degeneracy of the model that is broken as statistical power increases. Note that forcing a prior such that  $a_1 > 0$ , eliminating one of the peaks, also leads to essentially unchanged cosmology results (see below and in Sec. VII B).

It is worth bearing in mind that the  $a_1 = a_2 = 0$  point lies within the bounds of the 2D  $1\sigma$  contour, and it is plausible that the negative best-fit  $a_1$  is simply the result of a small true IA amplitude and noise (see also Appendix A). We also confirm that rerunning our fiducial analysis with a prior that forces  $a_1 > 0$  (in line with expectation from observations of red galaxies and the theory of tidal alignment) shifts the IA constraints, and results in a negative  $a_2$ , as consistent with our Y1 results, but does not appreciably alter the confidence contours in the  $S_8 - \Omega_m$  plane. If we restrict ourselves to the simpler (two-parameter) NLA model, the data still favor negative  $a_1$ , although at lower significance ( $a_1 = -0.09^{+0.20}_{-0.13}$ ). Again, this is lower than the Y1 results, both from cosmic shear alone and in combination with galaxy–galaxy lensing and galaxy clustering ( $a_1^{3\times 2\text{pt}} = 0.49^{+0.15}_{-0.15}$ ,  $a_1^{1\times 2\text{pt}} = 1.03^{+0.45}_{-0.57}$ ; see Dark Energy Survey Collaboration 2018 and Samuroff et al. 2019’s Table 5). The impact of the optimized  $\Lambda$ CDM analysis, which includes extra small-scale information, is to tighten the contours in the  $a_1 - a_2$  plane (compare the shaded green and black dashed lines in Fig. 8), and it does not qualitatively change our conclusions here.

One detail worth noting is that all of the posteriors presented in this paper include a contribution from shear ratios. The small-scale ( $\sim 2 - 6$  Mpc/h) galaxy–galaxy lensing information significantly improves our ability to constrain IAs (see Fig. 10 of Sánchez, Prat et al. 2021). That said, it has been demonstrated that the use of  $\gamma_i$  on large (instead of small) scales for the SR likelihood does not substantially change the favored IA scenario (Sánchez, Prat et al. 2021). It has also been shown that removing the SR likelihood altogether results in consistent, though broader, constraints in the  $a_1 - a_2$  plane (Amon et al. 2021).

We illustrate the redshift dependence of the inferred IA signal in Fig. 9, which can be compared to the analogous version from DES Y1 in Fig. 16 of Troxel et al. (2018), and also to Fig. 7 of Dark Energy Survey Collaboration (2021a). In IA models where we use a parametric redshift dependence (TATT and NLA), one can derive effective amplitudes for each redshift bin  $j$  as  $a_i^j = a_i (1 + \bar{z}_j / 1 + z_0)^{m_i}$ , where  $i \in (1, 2)$ . The points and error bars in Fig. 9 are the mean and marginalized 68% confidence contours of these derived parameters. In all cases the tidal alignment amplitude  $a_1$  is consistent between

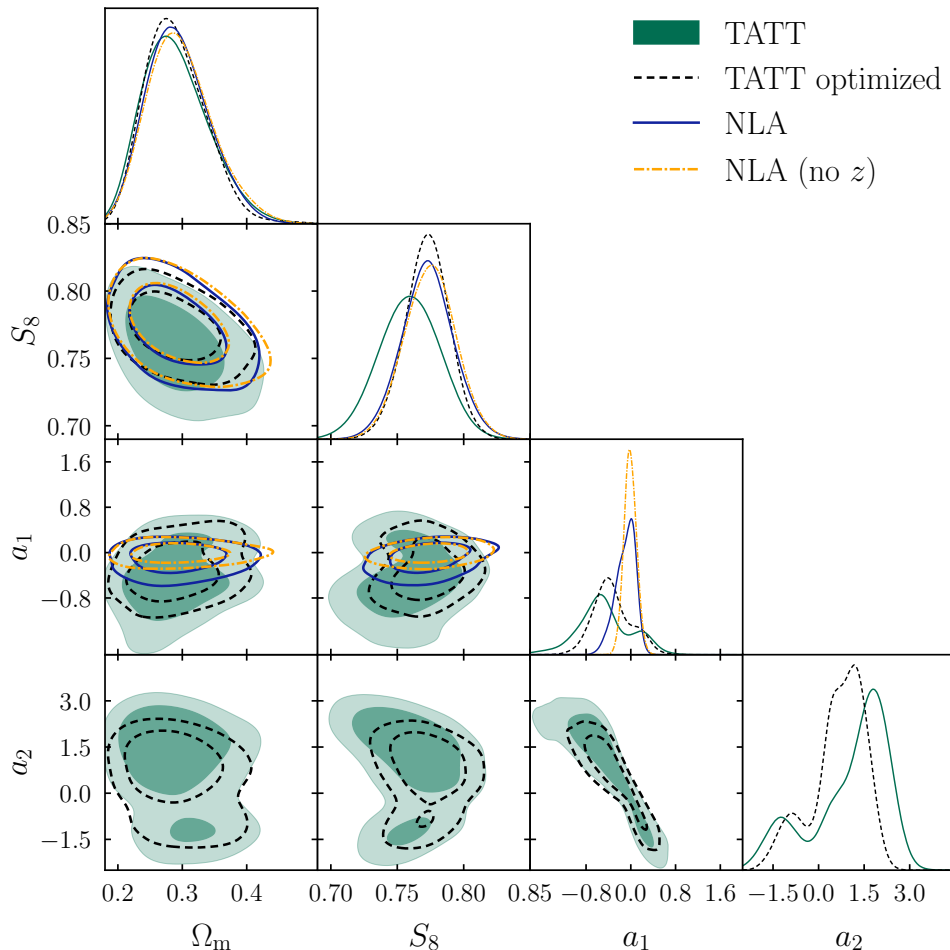


FIG. 8. Posterior parameter constraints from the fiducial  $\Lambda$ CDM cosmic shear analysis (TATT, green), which includes shear ratios as a fiducial choice. The solid blue (NLA) and dot-dashed yellow (NLA without redshift evolution) contours show the results of the same cosmic shear analysis, but using alternative (simpler) IA models. We also include our  $\Lambda$ CDM optimized TATT result (dashed black). As before, both 68% and 95% confidence levels are shown. We find that different IA models lead to consistent results and similar goodness-of-fit (see Table III), and the main difference between them is the change in constraining power.

models to  $\sim 1\sigma$ .

## 2. IA Model Selection

In addition to the basic results using variations of NLA and TATT in the IA implementation, we also perform a more rigorous model comparison in order to determine which IA model preferred by the data in a statistical sense. We explain this concept in more detail below.

First, we rank order models that are subspaces of TATT by their complexity and step up that list one at a time, usually by including a new parameter. At each step  $i$ , we re-run a chain to obtain the best-fit  $\chi^2$  and evidence of the model. We then compute the evidence ratios (see Sec. IV A) between the models at step  $i$  and  $i - 1$ , and between the model at step  $i$  and the fiducial TATT model. In doing this, we are in principle able to determine at which point in the complexity “ladder”

our data stops justifying the addition of extra IA parameters.

The order of the model complexity we use is not rigorously defined, but it follows the logic of the perturbative modeling approach. The simplest possible case is a model with no IA. We then consider tidal alignment (TA) less complex than models with tidal torquing (TT), and within these categories, models without redshift dependence are less complex. Thus, the simplest IA model under consideration has  $a_1$  free and all else held fixed, followed by  $a_1$  with a free power-law parameter for redshift evolution,  $\eta_1$ , and so on up to the (fiducial) 5-parameter TATT (see Table III).

In order to minimize sampling noise in the estimate of the best-fit  $\chi^2$ , we set POLYCHORD to output  $10\times$  more samples than its default. We verify that sampling the parameter space with a maximum posterior finder leads to  $\chi^2$  values that are insignificantly different from our best-fit ( $\Delta\chi^2 < 1$ ). We obtain the evidences as a standard output of POLYCHORD. At each step, we also report the constraints on the marginalized

IA parameters and their 68% confidence intervals.

With this method, we aim to find how many of the parameters meaningfully improve the fit to the data even though a model with an extra IA parameter might be *constrained*, the Occam factor in the evidence ratios can penalize that model if the required prior volume is excessive when compared to a simpler model, Eq. (7). We note that according to our analysis choices, this procedure can be carried out *after* the data and constraints are unblinded and so we could not have applied the same reasoning to unblinding (e.g. as a strategy to optimize the selection of a fiducial IA model).

We have performed this test with SR data vectors by using both DES Y3 lens samples (REDMAGIC and MAGLW) well as without including the SR likelihood at all, in case all information comes from  $\xi_{\pm}(\theta)$  on relatively large angular scales. For these three cases we find that the “simplest model” has the highest evidence, although it is only “marginally preferred” over the other models, based on the Jeffreys scale, reflecting the fact that our inferred IA amplitudes are consistent with zero. We find in all cases that, once the tidal torquing term  $a_2$  is free, the evidence ratios seem to suggest a preference for a redshift evolution. Apart from that, we find that the different SR lens samples yield slightly different levels of preference for the intermediate models such as NLA and NLA without  $z$ -evolution. Our results are shown in Table III, where stepping down line-by-line corresponds to increasing the IA model complexity, and for ease of interpretation we show only the case *without* the inclusion of the shear ratio likelihood (so note that the IA constraints presented in the table *do not exactly match the constraints reported elsewhere* in this work).

In summary, this test, as well those shown in Fig. 8, suggest that simpler IA models are a sufficient assumption for modeling the DES Y3 data. While this result is different from those in the pre-unblinding robustness tests which led to our choice of fiducial model (see Fig. 6), it is not inconsistent with expectations. Our earlier simulated tests used the best-fit TATT parameter values from the DES Y1 measurements. At these values, the impact of higher order IA contributions was large enough to require the full TATT model. However, the uncertainty in these Y1 measurements was fairly large, and they remain consistent with the overall smaller IA amplitudes found in this Y3 analysis. Smaller amplitudes allow for a less complex IA model, and indeed our current data appear to marginally prefer it.

Our results are consistent with other recent results, including Fortuna et al. (2021), Wright et al. (2020), which also suggest low IA amplitudes. While the tests presented here provide additional information to the community for planning future analyses, we emphasize that further study is needed on the underlying IA behavior of typical source galaxies and the interaction with other elements on the model, including baryonic feedback and photometric redshifts, as well as the impact of noise. We have physical reasons to believe that higher-order IA contributions exist at some level, and the lack of a significant detection here does not imply that these terms can be safely ignored at increased precision or for different source samples.

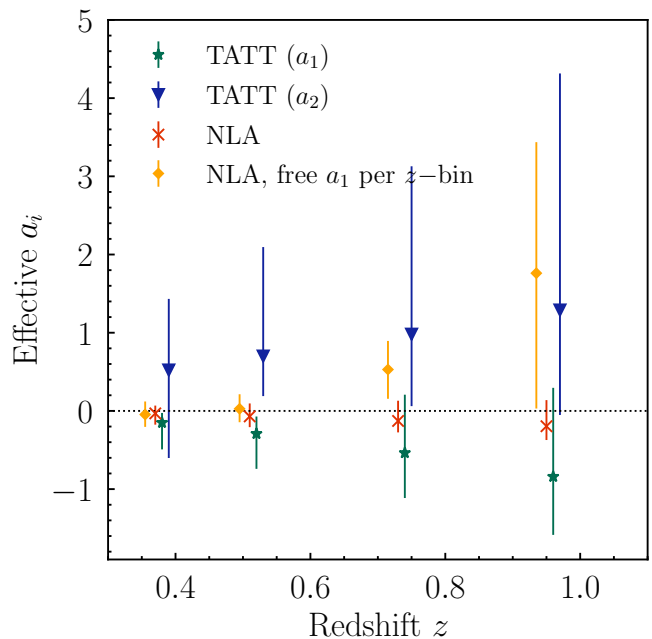


FIG. 9. Effective IA amplitudes  $a_{1,2}$  as a function of redshift. Note that the green and blue points (both labelled TATT) are the product of the same analysis, in which  $a_1$  and  $a_2$  were varied simultaneously. We report no clear evidence for redshift evolution in  $a_1$ , and relatively good agreement between models.

We encourage careful testing of IA model sufficiency in future analyses to avoid cosmological biases.

## VII. MODEL ROBUSTNESS TESTS ON REAL DATA

In this section we present a number of analysis permutations, with the aim of stress testing the results described in the previous section. The tests fall naturally into a number of groups, which are discussed in more detail in the following paragraphs. For an overall summary, see Fig. 10.

This exercise is distinct from the tests in Sec. V; whereas there we were seeking to validate our model implementation *prior to unblinding*, we are now seeking to test the robustness of our unblinded results to reasonable variations to the analysis choices. Our focus here is on various aspects of the theory modeling. For a complementary raft of internal-consistency and data oriented tests, we refer the reader to our companion paper (Amon et al. 2021). One can also find extensive catalog-level tests in Gatti, Sheldon et al. (2021), which includes a null detection of shear B-modes in the Y3 data using both COSEBIs and pseudo- $C_{\ell}$ s. Further tests of the PSF model and photo- $z$  catalogs are described in other Y3 papers (Jarvis et al. 2021; Buchs, Davis et al. 2019, Myles, Alarcon et al. 2020). Also note that the details of the Y3 methodology, including the tests described in this section, were chosen prior to unblinding, and are not informed by the results described in Sec. VI A.



IA Model (free parameters)	$\chi^2/\text{d.o.f}$	log Evidence	$R$ (w.r.t. TATT)	$R$ (w.r.t. above)	$a_1$	$\eta_1$	$a_2$	$\eta_2$	$b_{\text{TA}}$
No IAs	240.6 / 225	$3215.79 \pm 0.11$	$9.48 \pm 1.66$	N/A	-	-	-	-	-
NLA no $z$ -evo. ( $a_1$ )	238.6 / 224	$3213.89 \pm 0.12$	$1.42 \pm 0.30$	$0.15 \pm 0.03$	$0.34^{+0.25}_{-0.23}$	-	-	-	-
NLA ( $a_1, \eta_1$ )	238.3 / 224	$3214.07 \pm 0.13$	$1.70 \pm 0.36$	$1.19 \pm 0.24$	$0.36^{+0.43}_{-0.36}$	$1.66^{+3.26}_{-1.05}$	-	-	-
TA ( $a_1, \eta_1, b_{\text{TA}}$ )	238.8 / 224	$3213.87 \pm 0.13$	$1.38 \pm 0.25$	$0.81 \pm 0.14$	$0.27^{+0.35}_{-0.31}$	$2.10^{+2.89}_{-0.71}$	-	-	$0.83^{+0.31}_{-0.82}$
No $a_2$ $z$ -evo. ( $a_1, a_2, b_{\text{TA}}$ )	238.6 / 223	$3211.81 \pm 0.14$	$0.17 \pm 0.03$	$0.12 \pm 0.02$	$0.18^{+0.21}_{-0.30}$	-	$0.10^{+0.55}_{-0.57}$	-	$0.80^{+0.29}_{-0.78}$
No $a_2$ $z$ -evo. ( $a_1, \eta_1, a_2, b_{\text{TA}}$ )	238.2 / 223	$3212.09 \pm 0.14$	$0.23 \pm 0.04$	$1.32 \pm 0.26$	$-0.02^{+0.71}_{-0.31}$	$2.17^{+2.82}_{-0.70}$	$-0.27^{+0.59}_{-0.50}$	-	$0.87^{+0.38}_{-0.83}$
<b>TATT</b> ( $a_1, \eta_1, a_2, \eta_2, b_{\text{TA}}$ )	233.1 / 222	$3213.54 \pm 0.13$	1	$4.28 \pm 0.83$	$-0.24^{+0.98}_{-0.41}$	$2.38^{+2.62}_{-0.61}$	$0.63^{+1.93}_{-1.89}$	$3.11^{+1.77}_{-0.31}$	$0.87^{+0.38}_{-0.84}$

TABLE III: Marginalized constraints on IA parameters, best-fit  $\chi^2$  and evidence ratio  $R$  metrics for IA models of increasing complexity on  $\xi_{\pm}(\theta)$  *without including the shear ratio likelihood*. In each row, the evidence ratios assume TATT & the model in the row above are in the denominator (so  $R > 1$  implies preference for the model defined in the corresponding row).

Marginalized constraints are defined here as the mean of posteriors and their 68% uncertainties. Empty values in the table correspond to parameters that are fixed to zero.

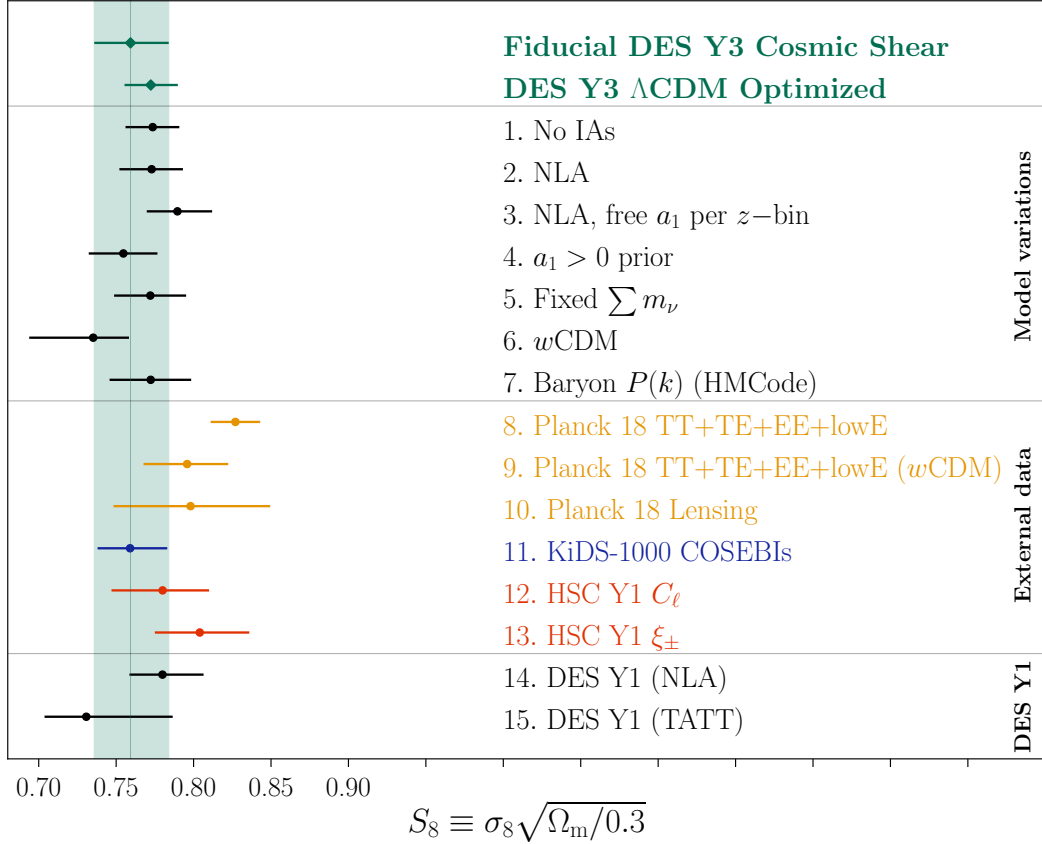


FIG. 10. A summary of marginalized 1D constraints on  $S_8$  shown in this paper. In the upper-most panel we show the main cosmological results of this paper: the DES Y3 cosmic shear constraints using fiducial and  $\Lambda$ CDM Optimized scale cuts. The panel below (marked “model variations”, rows 1-7) shows a range of modified analyses, designed to test the robustness of the fiducial result, which are detailed in Sec. VI & VII). In the lower two panels we show equivalent constraints on  $S_8$ , both from external data (rows 8-13), and DES Y1 (rows 14-15). Points and error bars correspond to the marginal posterior mean and 68% confidence interval on  $S_8$ , with the exception of KiDS & HSC (rows 11-13) for which we report nominal headline results. Rows 1-10 are obtained using the DES Y3 fiducial analysis choices (including cosmology priors), while external lensing and DES Y1 rows 11-15 are *not re-processed* to match exactly the Y3 model, prior and scale cuts.

### A. Baryonic $P(k)$ & Neutrinos

One significant source of systematic uncertainty is in the modeling of the small- to intermediate-scale matter power spectrum. Uncertainties at high  $k$  arise both as a result of nonlinear clustering, a process that is typically calibrated using  $N$ -body simulations, and baryonic physics. Although we have verified our insensitivity to reasonable changes in baryonic and nonlinear growth scenarios in Sec. IV G, the tests there use noiseless synthetic data.

In order to assess the dependence of our results on our particular choice of  $P_\delta(k)$  model, we repeat our fiducial  $\Lambda$ CDM analysis, with an alternative power spectrum estimator. That is, instead of HALOFIT, which is our baseline choice, we use the halo model of Mead et al. (2015), which has a nominal accuracy of 5% at  $k < 10h\text{Mpc}^{-1}$ , as assessed by comparison with the Coyote Universe simulations. The halo model includes baryonic effects via two additional parameters: an amplitude  $B$  which governs the halo concentration-mass relation, and  $\eta_0$ , which is referred to as the halo bloating factor. We marginalize over both with wide flat priors  $\eta_0 = [0.4, 1]$ ,  $B = [1, 7.5]$ . These priors are wider than those utilized by, e.g., Asgari et al. (2021) and are intentionally chosen that way so we are agnostic with respect to the baryonic effects tested in Mead et al. (2015). Another difference with respect to the analysis of Asgari et al. (2021) is that we free both parameters independently, as opposed to assuming a linear relation between them.

The results of this exercise are shown in Fig. 10 (labelled ‘‘HMCode  $P(k)$ ’’) and in Fig. 11. As shown there, there is a small  $0.5\sigma$  shift in  $S_8$  due to this substitution. We do not, however, interpret this as the correction of residual baryonic physics present in the data, but rather a result of projection effects. Indeed, we verify that the halo-model parameters  $B, \eta_0$  are completely unconstrained by the data within the prior bounds; this is a consequence of our conservative scale cuts, which are designed to remove angular scales on which baryonic feedback processes (and so the halo-model parameters) enter. The small shift is, therefore, simply an artifact of projecting the high-dimensional posteriors onto the 1D plane. To support this point, we verify that shifts in the same direction appear when analysing a gravity-only synthetic data vector with the same wide baryon priors. We see no significant improvement in the goodness-of-fit of our data when employing HMCODE (a  $\chi^2$  of 236.8 vs. our fiducial 237.7, with 2 added parameters).

In addition to the tests described above, we also repeat our fiducial analysis with the neutrino mass density parameter  $\Omega_\nu h^2$  fixed. For this exercise, we set the sum of the neutrino masses,  $\sum m_\nu$ , to the lower limit obtained by oscillation experiments, assuming the normal (non-inverted) hierarchy, of 0.06eV (Capozzi et al. 2014, Esteban et al. 2019). We maintain our fiducial choice of 3 degenerate neutrino species (Krause et al. 2021). The neutrino mass parameter is not constrained by cosmic shear alone, so the small upward shift of around  $\sim 0.2\sigma$  we observe in the 1D marginal  $S_8$  constraint shown in Fig. 10 is again thought to be an effect of the projection of the high-dimensional parameter space with a different prior volume.

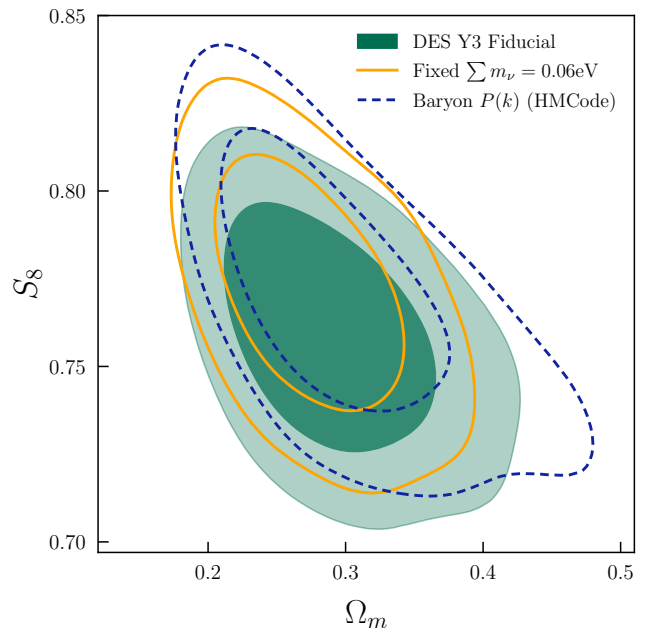


FIG. 11. The marginal posterior distribution in the  $S_8 \times \Omega_m$  plane for the model variations on the baryonic power spectrum (HMCode with 2 free parameters) and neutrino mass density (fixed at minimum mass). The results of our fiducial analysis is also shown (green). We observe consistency between these cases and our fiducial analysis, and do not interpret shifts as pointing to insufficiency in the analysis (see text).

We show the 2D parameter constraints in the  $S_8$ - $\Omega_m$  plane under the variations of assumptions in the baryonic power spectrum and the neutrino mass density in Fig. 11. We note that despite the small differences in marginalized  $S_8$  constraints, these analysis variations are still fully consistent with our fiducial analysis.

### B. Intrinsic Alignments

In this section we explore how plausible variations in our IA modeling can propagate to cosmology. In addition to the fiducial TATT model, we consider a number of alternatives, which are detailed below (see also Fig. 10). The variations cover a range of complexity scenarios, from the simplest case of null IAs (i.e all parameters fixed to zero), through NLA and TATT to a version with added flexibility in redshift.

In Fig. 8 we consider our baseline result (green), alongside the equivalent using the two-parameter NLA model (blue; Troxel et al. 2018 and Hikage et al. 2019’s fiducial choice), and an even simpler one-parameter version with no redshift dependence (dot-dashed yellow; Asgari et al. 2021’s fiducial choice). Although only at the level of  $\sim 0.5\sigma$ , we see a shift in  $S_8$  when switching between NLA and TATT, which is roughly consistent with the findings of Samuroff et al. (2019) in DES Y1. This shift is not accompanied by a large improvement in the  $\chi^2$  per degree of freedom, but there is some preference for

the simpler mode in terms of Bayesian evidence (see Sec. VIC and Table III). It is also notable that the variants of NLA with and without redshift variation give virtually identical cosmology results, primarily because the extra  $\eta_1$  parameter is poorly constrained, and relatively uncorrelated with  $S_8$ . We also to note briefly that the impact of switching IA models also slightly rotates the best constrained direction in parameter space, with  $\alpha = 0.576$  for NLA and  $\alpha = 0.586$  in the fiducial TATT, when one defines the lensing amplitude as  $\Sigma_8 = \sigma_8(\Omega_m/0.3)^\alpha$ .

Given the negative  $a_1$  and hints of bimodality seen in Fig. 8, we run a version of the Y3 cosmic shear analysis with a restrictive prior  $a_1 = [0, 5]$ .<sup>15</sup> This results in almost no change to the best-fit  $S_8$  (compare lines 1 and 6 in Fig. 10), although it does tighten the error bar slightly by restricting the posterior (row 6 of Fig. 10), which implies that the fact that our fiducial results encompass a region of negative IA space is not driving our cosmological constraints in a particular direction.

We also present a case with a flexible version of the NLA model, shown in Fig. 10 (line 5, labelled “NLA, free  $a_1$  per  $z$ -bin”). The basic idea here is to test how limiting our assumption of power law redshift dependence is in this context, since we have no first-principles reason to expect IAs should obey this particular scaling. Interestingly, adding flexibility via extra TATT parameters, and via the ability to deviate from a power law redshift evolution, is seen to push  $S_8$  in opposite directions relative to the basic NLA (c.f. lines 1, 4 and 5 in Fig. 10). While we find a shift of around  $1\sigma$  w.r.t TATT ( $0.5\sigma$  w.r.t to NLA), the per-bin NLA model is significantly disfavored. The Bayesian evidence ratios strongly favor TATT and NLA over the per-bin NLA model (with  $R = 237 \pm 60$  and  $202 \pm 49$  respectively), and in the latter model we additionally see a significantly worse  $\chi^2$  (increased by about 10 with respect to TATT while the effective number of degrees of freedom is decreased by 1).

Taken as a whole, the tests described above paint a consistent picture. Apart from NLA per-bin, the IA modifications we explored cause shifts no greater than  $\sim 0.5\sigma$  in  $S_8$ , even in the most extreme scenario where we neglect IAs entirely. In NLA per-bin, we find a shift of  $1\sigma$  with respect to our fiducial constraints, but also find that this model is disfavored by evidence ratios and goodness-of-fit tests. We therefore do not believe we that our IA modeling is *insufficient* to describe the data and, through the IA model selection described in Sec. VIC2, actually have evidence that it might be simplified in future analyses. We do find a change in constraining power between NLA and TATT, with NLA (at fiducial scale cuts) being comparable to TATT at the  $\Lambda$ CDM-Optimized scale cuts. This suggests that our uncertainty budget is dominated by the lack of knowledge in modeling astrophysics on small cosmological scales. This is verified by our companion Amon et al. (2021), Sec. XII.

<sup>15</sup> Our baseline prior bounds  $a_1 = [-5, 5]$  were designed to be uninformative and avoid prior edge effects. It is, however, reasonable to expect that  $a_1 > 0$ , in the absence of systematics.

## VIII. COMPARISON WITH EXTERNAL DATA

After having verified that DES-only results are robust with respect to changes in astrophysics modeling assumptions, we now seek to place our results in the context of the wider field and to quantify tension with respect to subsets of external data.

### A. External Data Sets

We describe a number of external data sets (see Table VIII A), to which we will compare our results. Where appropriate, we recompute the cosmological posteriors in order to facilitate a meaningful comparison. The marginalized  $S_8$  constraints from the external data sets that meaningfully constrain this parameter alone are shown in the lower half of Fig. 10. The data sets we consider in this paper are largely common to those described in Dark Energy Survey Collaboration (2018), Dark Energy Survey Collaboration (2019) and Troxel et al. (2018), with some more recent updates. Among these, as described in Sec. 1, the KiDS and HSC surveys are Stage-III WL surveys like DES. The full set of surveys we consider is:

- **KiDS-1000:** The KiDS weak lensing data comprise roughly 1000 square degrees and 21M galaxies ( $n_{\text{eff}} \sim 6.2 / \text{arcmin}^2$ ), and are described in Giblin et al. (2021). In their latest results papers, (Asgari et al. 2021) present cosmic shear analyses using three different statistics; since they clearly designate their COSEBIs-based results as their fiducial analysis, we compare with these here. Note that we do not recompute the posteriors, but rather compare with the published results. In particular, one should be aware that our choice of IA model, nonlinear power spectrum and cosmological priors all differ from the fiducial analysis of Asgari et al. (2021).
- **HSC Y1:** The first year HSC lensing data are drawn from 137 square degrees, but go significantly deeper than either KiDS or DES, reaching  $n_{\text{eff}} \sim 22 / \text{arcmin}^2$ . The data are described in Mandelbaum et al. (2018a), and are calibrated using image simulations (Mandelbaum et al. 2018b). As in the case of KiDS, we do not reanalyze the two-point data, instead comparing with the published results. Since, at the time of writing, there is no reason to prefer one over the other, we show both the real space analysis of Hamana et al. (2020) and the power spectrum analysis of Hikage et al. (2019) in Fig. 10. The differences between the two are thought to be statistical, due to different Fourier mode sensitivities. For the sake of clarity, we choose to show the latter in the visual comparison of Fig. 13.
- **eBOSS:** We include spectroscopic baryon acoustic oscillation (BAO) measurements from eBOSS (Alam et al. 2021). The BAO likelihood is assumed to be independent of DES, but we do recompute the posterior in our choice of cosmological parameter space (i.e. sampling  $A_s$  and  $\Omega_m$ , and with the sum of the neutrino masses free).

- **Pantheon Supernovae:** We also include the luminosity distances from Type Ia supernovae from the Pantheon sample (Scolnic et al. 2018); 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 Main:** These data are the final release from the Planck Cosmic Microwave Background (CMB) experiment (Planck Collaboration 2020a). We incorporate the primary  $TT$  data on scales  $30 < \ell < 2508$ , and also the joint temperature and polarization measurements ( $TT + TE + EE + BB$ ) at  $2 < \ell < 30$ . As in previous analyses, we recompute the CMB likelihood in our fiducial parameter space, including neutrinos.
- **Planck 2018 Lensing:** We also consider CMB lensing from the Planck survey (Planck Collaboration 2020c) as a separate data set. These data probe an intermediate redshift  $z \sim 2$ , which is slightly higher than DES and somewhat lower than CMB temperature and polarization.

## B. Quantifying Tension

Assuring that the data collected by two different experiments have a quantified degree of agreement is crucial when performing a joint probes analysis. Therefore, for DES Y3, we have presented an thorough study on the ability of different tension metrics to identify inconsistencies amongst cosmological parameters measured by different experiments. We used simulated data to compute the predictions of different tension metrics belonging to two classes: Evidence-based methods and parameter-space methods. A robust way to quantify possible tensions is by combining those two types of metrics, as they answer somewhat different, but yet complementary, questions (Lemos et al. 2020).

In DES, we utilize priors that are deliberately wide and uninformative such that "DES-only" constraints can be obtained. In this case, assessing tension utilizing the Bayes Ratio, a broadly used evidence-based tension metric, can produce misleading results since it is largely dependent on the prior volume as discussed in Handley & Lemos (2019b). To avoid this problem, for our choice of an evidence-based metric we compute the Bayesian Suspiciousness (Handley & Lemos 2019a) instead, a method that corrects for the prior dependence.

Consider two independent data sets A and B. The motivation behind Suspiciousness is that the Bayes Ratio  $R$  can be divided into two parts: the first one captures the prior dependence, i.e. the probability of the data sets matching given the prior width, which is quantified by the information ratio  $I$ :

$$\log I \equiv \mathcal{D}_A + \mathcal{D}_B - \mathcal{D}_{AB}, \quad (31)$$

where

$$\mathcal{D}_D \equiv \int \mathcal{P}_D \log \left( \frac{\mathcal{P}_D}{\Pi} \right) d\theta, \quad (32)$$

is the Kullback–Leibler Divergence (Kullback & Leibler 1951), that can be understood as the amount of information that has been gained going from the prior  $\Pi$  to the posterior  $\mathcal{P}$ . The lower index  $D$  denotes the data set from which the posterior is derived (A, B, or the joint data vector AB). The second part is the so-called Bayesian Suspiciousness  $S$ , which is the part of the Bayes Ratio left after subtracting the dependence on the prior, leaving only on the actual differences between the posteriors:

$$\log S = \log R - \log I. \quad (33)$$

All the quantities required to compute the Suspiciousness metric are provided by a single nested sampling chain, just as the ones required to compute the Bayes Ratio are, meaning the computational cost is the same in both cases. The necessary tools are implemented in the python package `anesthetic`<sup>16</sup> (Handley 2019).

In addition, we calculate a Monte Carlo estimate of the probability of a parameter difference, set out in Raveri et al. (2020), as our parameter space-based method. It relies on the calculation of the parameter difference probability density  $\mathcal{P}(\Delta\theta)$  which, in the case of two uncorrelated data sets, is simply the convolution integral:

$$\mathcal{P}(\Delta\theta) = \int_{V_p} \mathcal{P}_A(\theta) \mathcal{P}_B(\theta - \Delta\theta) d\theta, \quad (34)$$

where  $V_p$  is the prior support and  $\mathcal{P}_A$  and  $\mathcal{P}_B$  are the two posterior distributions of parameters. The probability of an actual shift in parameter space is obtained from the density of parameter shifts:

$$\Delta = \int_{\mathcal{P}(\Delta\theta) > \mathcal{P}(0)} \mathcal{P}(\Delta\theta) d\Delta\theta, \quad (35)$$

which is the posterior mass above the contour of constant probability for no shift,  $\Delta\theta = 0$ .

Usually we only have discrete representations of the posterior samples and, since we are working with high dimensional parameter spaces, the posterior samples cannot be easily interpolated to obtain a continuous function. Therefore, the integral in Eq. (35) needs to be performed using a Monte Carlo approach.

The idea is to compute the Kernel Density Estimate (KDE) probability that  $\Delta\theta = 0$  and the same at each of the samples from the parameter difference posterior. Then, the estimate of the integral in Eq. (35) is given by the number of samples with non-zero KDE probability, divided by the total number of samples. These computations are done using the `tensiometer`<sup>17</sup> code.

<sup>16</sup> <https://github.com/williamjameshandley/anesthetic>

<sup>17</sup> <https://github.com/mraveri/tensiometer>

Data set	Type	Median Redshift	Area [sq. deg]	Source
DES Y3	WL	0.62	4143	This work
KiDS-1000	WL	0.54	777	Asgari et al. (2021)
HSC Y1	WL	0.81	137	Hamana et al. (2020), Hikage et al. (2019)
eBOSS	BAO	0.7	6813	Alam et al. (2021)
Pantheon	SNe	1.0	-	Scolnic et al. (2018)
Planck Main	CMB $TT + TE + EE + \text{lowE}$	1090	full-sky	Planck Collaboration (2020a)
Planck Lensing	CMB Lensing	2.0	full-sky	Planck Collaboration (2020c)

TABLE IV: A summary of the external data sets used in this paper. More details can be found in Sec. VIII and the references listed.

### 1. Compatibility between Low- $z$ and High- $z$ Data

We now assess the consistency between high- and low-redshift cosmological probes. As in Dark Energy Survey Collaboration (2021a), we conclude that two data sets are statistically consistent if the  $p$ -value implied by our tension metrics is greater than 0.01. This same standard value was utilized in our internal consistency tests in Amon et al. (2021). In order to quote tension metrics, we first recompute the low- $z$  (here defined as the combination BAO + SNe) and Planck 2018 posteriors with priors on cosmological parameters matched to those of DES Y3, as defined in Table I. In addition to assessing tension relative to our fiducial Y3 analysis, we also explore whether switching to the  $\Lambda$ CDM-Optimized scale cuts significantly changes our findings, since the one-dimensional result in Fig. 10 can offer hints, but not quantify the agreement between data sets.

Fig. 12 shows the posteriors from DES as well as low- $z$  and Planck 2018 in  $\Lambda$ CDM. In the full parameter space, for DES Y3 cosmic shear *versus* BAO+SNe in  $\Lambda$ CDM we find:

$$\begin{aligned} \text{Suspiciousness:} & \quad 0.5\sigma \pm 0.3\sigma \quad (\text{DES vs Low-}z \text{ } \Lambda\text{CDM}) \\ \text{Parameter shift:} & \quad 0.4\sigma \quad (\text{DES vs Low-}z \text{ } \Lambda\text{CDM}), \end{aligned}$$

which implies  $p(0.5\sigma) = 0.62 > 0.01$ , and we conclude these data sets are consistent. This confirms more rigorously what is seen qualitatively in Fig. 12.

We similarly quantify the compatibility between DES Y3 cosmic shear and Planck 2018 in  $\Lambda$ CDM:

$$\begin{aligned} \text{Suspiciousness:} & \quad 2.3\sigma \pm 0.3\sigma \quad (\text{DES vs Planck } \Lambda\text{CDM}) \\ \text{Parameter shift:} & \quad 2.3\sigma \quad (\text{DES vs Planck } \Lambda\text{CDM}), \end{aligned}$$

with a  $p$ -value  $p(2.3\sigma) = 0.02 > 0.01$ . We thus find that, according to our tension metrics, these data sets are consistent. We additionally explore the consistency of DES Y3 cosmic shear with Planck 2018 in our  $\Lambda$ CDM-Optimized setup and find:

$$\begin{aligned} \text{Suspiciousness:} & \quad 2.0\sigma \pm 0.4\sigma \quad (\text{DES optim. vs Planck } \Lambda\text{CDM}) \\ \text{Parameter shift:} & \quad 2.1\sigma \quad (\text{DES optim. vs Planck } \Lambda\text{CDM}), \end{aligned}$$

which yields approximately  $p(2.0\sigma) = 0.05 > 0.01$ , again implying that these data sets are consistent. That result can be expected, at least qualitatively, by inspecting Figs. 10 and 12: while the constraint on  $S_8$  is improved in the optimized analysis, there is also a small shift towards Planck.

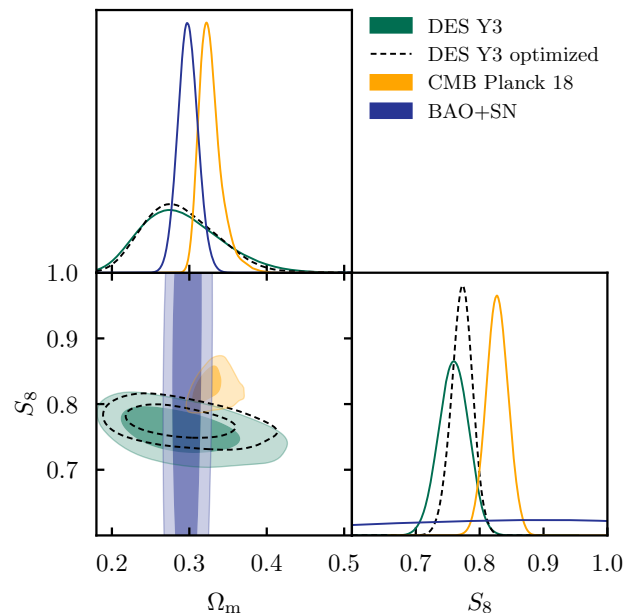


FIG. 12. DES Y3 and external data constraints from low and high redshift probes in  $\Lambda$ CDM. We present our fiducial and  $\Lambda$ CDM optimized constraints (green and black-dashed) in comparison with Planck 2018 (TT+EE+TE+lowE, no lensing; yellow) and the combination of BAO and type Ia SNe (SDSS BOSS and Pantheon respectively; blue). In all cases we show both the 68% and 95% confidence limits. We find no evidence for statistical inconsistency between DES Y3 cosmic shear and either external data set.

### 2. Comparing Weak Lensing Surveys

We show our fiducial DES Y3  $\Lambda$ CDM results (green shaded), alongside those of a number of contemporary weak lensing surveys in Fig. 13. Also shown (yellow) are the equivalent constraints from most recent Planck CMB data release (without lensing). Most notably, all the published cosmic shear analyses (including earlier, less constraining, results that are not shown here) favor lower  $S_8$  than Planck, to varying degrees of significance. Although the cosmic shear surveys are independent, both in the sense that there is limited overlap between the catalogues (see Fig. 1), and that the measurement and analysis pipelines are largely separate, it is remarkable that

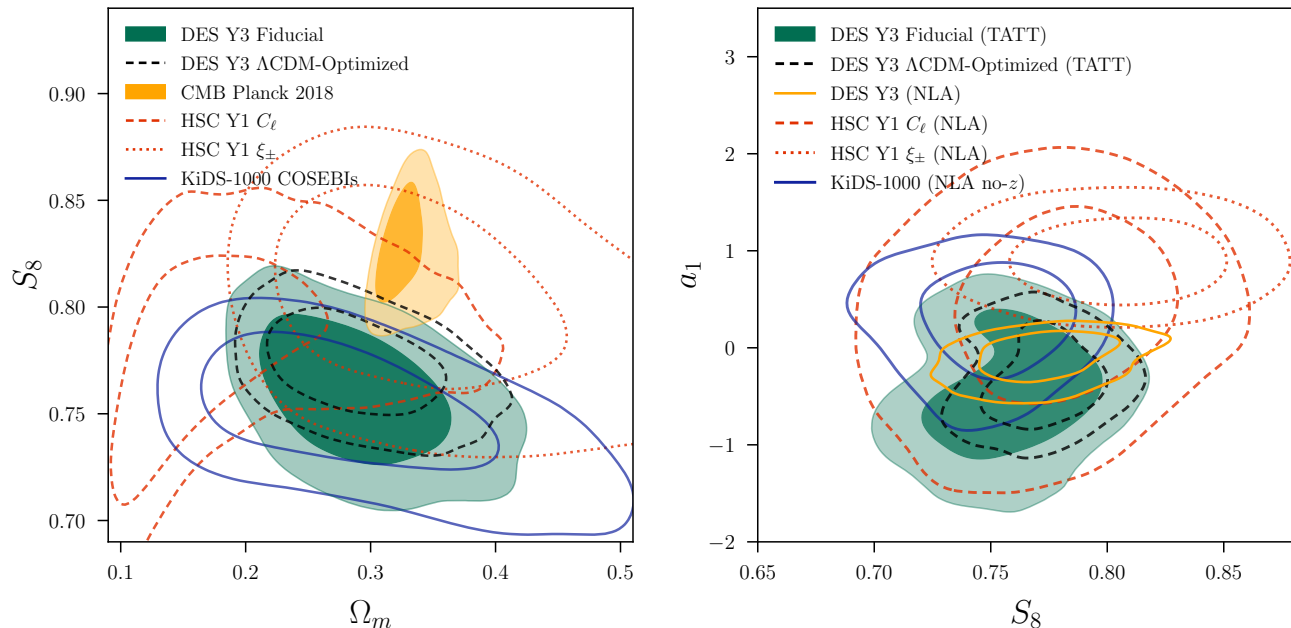


FIG. 13. Posterior constraints on cosmology (left panel) and intrinsic alignment parameters (right panel) from contemporary weak lensing surveys. With the exception of Planck 2018, which is re-analyzed with DES Y3 priors, the weak lensing posterior data above are plotted as published by each collaboration, therefore a direct, quantitative comparison is complicated due to differing parameter priors, modeling and calibration choices. An overall qualitative agreement is seen in both panels, within approximately  $1\sigma$ , with lensing surveys predicting similarly low  $S_8$  and consistent  $a_1$  amplitude.

their  $S_8 \times \Omega_m$  posteriors overlap significantly, especially given that these weak lensing analyses are carried out blind.

That said, the naive comparison of the surveys is complicated by the fact that there are significant differences in the analysis choices underlying the published results, and unifying them becomes a crucial task (Chang et al. 2019). Though it is quite possible to compute the metrics discussed in the previous section, interpreting them would be complicated, as it would be difficult to determine if any apparent tension were real, or due to differences in priors (see the discussion in Lemos et al. 2020). A proper quantitative comparison would require matching priors and other analysis choices, and would ideally involve input from the different collaborations; such an exercise would be in line with the recommendations of Chang et al. 2019 and is left for future work.

We can, however, compare our findings with the tension analyses that the various collaborations choose to present. For example, in DES Y1 cosmic shear the discrepancy relative to Planck 2018 was assessed to be at the level of  $1.0\sigma$  and  $0.5 \pm 0.3\sigma$  using the parameter shift and Suspiciousness metrics respectively (Lemos et al. 2020). These numbers can be compared directly with our findings in Sec. VIII B 1. We see that the discrepancy between cosmic shear and Planck 2018 has increased substantially in DES Y3 in comparison with DES Y1, from  $1.0\sigma$  to  $2.3\sigma$ . The latter number is, coincidentally, the same difference found in the comparison between Planck and the  $3 \times 2$ pt analysis of DES Y1 (including galaxy clustering and galaxy-galaxy lensing).

The KiDS-1000 cosmic shear analysis of Asgari et al. (2021)

opted to quantify their agreement with Planck using the Bayes ratio. They find “substantial” evidence for disagreement, with the caveat that the performance of this metric is known to be prior dependent. Though they also compute a Suspiciousness, they do not quote an interpretable value in terms of  $\sigma$ , due to difficulties in computing effective dimensionalities (see their App. B3). The approach of HSC differs slightly between their two published analyses. Hamana et al. 2020 (their Sec 6.6) rely on a comparison of projected contours, reporting no tension due to the apparent overlap in the  $\sigma_8 \times \Omega_m$  plane. On the other hand, Hikage et al. (2019) employ both the Bayes ratio and also differences in the log-likelihood at the MAP point in parameter space (see their Sec 6.3 and Raveri & Hu 2019b). They report no evidence for inconsistency based on these metrics.

In the right-hand panel of Fig. 13 we also show the IA model constraints from the various surveys. Although they are roughly consistent (to within  $\sim 1\sigma$ ), it is worth being careful here. In all cases but KiDS-1000, the IA model has more flexibility than the one amplitude (an additional redshift scaling in HSC and DES Y1, and the extra TATT parameters in Y3), which means, for a given  $a_1$ , the predicted IA signal is not necessarily identical between the surveys. It is also true that, unlike cosmological parameters, IAs are *expected* to be dependent on both the galaxy selection and the shape measurement method. Differences, while interesting from the modeling perspective, are not necessarily a cause for concern.

## IX. CONCLUSIONS

This paper and its companion, [Amon et al. \(2021\)](#), present together the cosmological constraints from cosmic shear with over  $\sim 100$  million galaxies from Dark Energy Survey Year 3 (DES Y3) shape catalogs, which cover 4143 square degrees. We model the cosmic shear signal in  $\Lambda$ CDM and constrain the lensing amplitude with 3% precision, finding  $S_8 = 0.759^{+0.023}_{-0.025}$ . Our best constraint on  $\Lambda$ CDM is of  $S_8 = 0.772^{+0.018}_{-0.017}$ , at 2% precision, when optimizing the angular scales while still maintaining biases under control. Our results qualitatively agree well with previous lensing analyses based on the KiDS-1000, DES Y1 and HSC data sets, all of which favor lower  $S_8$  than the most recent CMB measurements. Using quantitative tension metrics based on the full parameter space, we report a  $2.3\sigma$  difference between our DES Y3 results and Planck 2018 (a  $p$ -value of 0.02), and we consider these data sets to be statistically consistent. At our present constraining power, we do not report a meaningful constraint on the equation-of-state parameter of dark energy in  $w$ CDM. That is, the cosmic shear likelihood does not significantly add information about the parameter  $w$  beyond its prior.

Throughout this paper, we have focused on the aspects of the DES Y3 cosmic shear analysis relating to astrophysical modeling. The overarching philosophy that guided this work and its conclusions can be summarized as follows. Firstly, in a blinded fashion, we test our modeling assumptions on synthetic data obtained from theory pipelines and N-body mocks, in order to validate our fiducial analysis choices. We then obtain cosmology constraints from shear data that are verified to be internally consistent and calibrated to high confidence by our companion paper ([Amon et al. 2021](#)). Finally, we proceed to relax our fiducial theory assumptions and vary model parameterizations, finding that the results to be robust to these variations, and then compare our constraints with those of external probes. The main conclusions of this paper are:

- Informed by previous intrinsic alignment (IA) studies in DES Y1 data, we account for IAs with a 5-parameter model that includes tidal alignments and tidal torquing (TATT). We present tests of this modeling assumption on analytically generated data as well as realistic MICE (N-body) simulations, and find that TATT is able to capture complex IA signals, as well as that of simpler models such as NLA ([Fig. 6](#));
- We select angular scales conservatively to mitigate the effect of baryonic physics, the dominant small-scale systematic for this analysis. We show that our scale cuts suppress baryonic contamination as inferred from hydrodynamic simulations with varying levels of feedback strength, and utilize a gravity-only model for the matter power spectrum. We verify in synthetic data that residual biases due to baryons are well below  $0.3\sigma$  in the  $\Omega_m \times S_8$  plane, and that our selected scales are insensitive to uncertainties in the nonlinear power spectrum and in higher order shear contributions ([Fig. 5](#));
- The tension metrics we use in the full parameter space

yield a  $2.3\sigma$  discrepancy with respect to Planck (2018), with the leading contribution to the tension being our lower  $S_8$ . External probes at low- $z$  (BAO+SNe), which are sensitive to  $\Omega_m$ , are within  $0.5\sigma$  under the same metrics. We regard both these probes as statistically consistent with DES cosmic shear. We also find qualitative agreement between our data and external lensing results from HSC Y1 and KiDS-1000, all of which yield lower nominal  $S_8$  values in comparison with Planck 2018 ([Figs. 7, 12](#) and left panel of [13](#));

- We demonstrate that our posteriors on IA and cosmological parameters are consistent within  $1\sigma$  as we vary the parameterization of intrinsic alignments by simplifying our fiducial model from TATT to NLA with free/fixed redshift dependence. This is also true when we allow for free  $a_1^i$  amplitudes on each redshift bin  $i$ . We demonstrate that our posteriors on cosmological parameters are stable when inference is carried out at fixed neutrino mass and with free baryonic feedback parameters ([Figs. 9, 10](#) and [11](#));
- We perform a detailed Bayesian evidence-based model selection for intrinsic alignments, finding that our data shows a weak preference for simpler (and better constrained) parameterizations. We additionally find that the best-fit IA amplitudes in DES Y3 are smaller than those in Y1. However, the two are consistent with each other, and also with other lensing surveys, although quantitative statements are challenging without further studies of e.g. sample selections and model differences. In combination with the fact that posteriors on  $S_8 \times \Omega_m$  from simpler IA models are consistent with our fiducial choice (TATT), these findings point to less complex parameterizations, such as NLA, being a sufficient and unbiased description of our data ([Fig. 8](#), [Table III](#) and right panel of [Fig. 13](#)).

Additionally, we share a summary of the main conclusions of our companion paper ([Amon et al. 2021](#)), demonstrated with the same Y3 cosmic shear analysis, and point the reader to that paper for details:

- The analysis is shown to be robust to the choice of redshift calibration sample, either photometric or spectroscopic, methodology and the modeling of redshift uncertainty, within  $\sim 0.5\sigma$ .
- We model the impact of blending using state-of-the-art image simulations and show that our posteriors are stable to this correction, within  $0.5\sigma$ .
- The analysis passes all internal consistency tests, finding that the cosmology is stable across redshifts, angular scales and measurement statistics.
- The impact of additive shear systematics, such as PSF contamination and B-modes is assessed and found to be negligible for the analysis.

- We investigate the limiting factors of the cosmic shear constraints and find observational systematics are sub-dominant compared to systematics due to modeling astrophysical effects.

The results presented in this study show that cosmic shear has reached the requirements of percent level precision: every source of systematic uncertainty must be controlled to better than a percent of the signal, since our statistical uncertainty on the amplitude parameter  $S_8$  is three percent. When the first detections of cosmic shear were published in 2000, it was far from a given that this level of precision was achievable. With the full DES survey awaiting analysis, and the much larger surveys from Euclid and the Rubin Observatory’s LSST starting in a few years, it is interesting to consider prospects for further qualitative advances.

On the theoretical side, we are limited by our ability to describe the astrophysics of small scales, in particular baryons and IA. Baryonic physics is the driver of our scale cuts, and we will need to rely on advances in hydrodynamic simulations, constrained by complementary observations such as Sunyaev-Zel’dovich (SZ) maps, otherwise the necessary cuts and/or marginalization will cause more information to be discarded. Conversely, we have found that intrinsic alignments are likely to contribute less than the most conservative forecasts, but more studies are required to extend our findings to arbitrary galaxy samples. Parallel advances in shear estimation and calibration and the photo- $z$ s of source galaxies are required to keep pace with statistical errors. Beyond cosmic shear, a suite of statistical measures that capture the non-Gaussian information in the shear field have been developed, such as peak statistics and three-point correlations. These will add complementary information and also serve as cross-checks on the robustness of the signal. We can look forward to new cosmological tests with the application of these approaches to DES and future surveys.

Finally, we reiterate that, while the assumption of  $\Lambda$ CDM as the ultimate end-to-end model connecting the early- and late- Universe has withstood another test, our result should be understood within a broader context. It is still an open puzzle that modern weak lensing surveys, independently and in blind analyses, find a lower lensing amplitude than predicted by the CMB, and the difference between DES cosmic shear itself with respect to Planck has increased from  $1.0\sigma$  in DES Y1 to  $2.3\sigma$  in Y3. More effort is required to quantify the agreement between lensing studies, especially towards unifying and homogenizing their analyses and exploiting their complementarities.

## ACKNOWLEDGEMENTS

LFS and BJ were supported in part by the US Department of Energy grant DE-SC0007901. We thank Rachel Mandelbaum for many useful discussions during the writing of this paper. We would like to kindly ask the readers to cite this manuscript with dual-authorship as Secco & Samuroff *et al.* whenever possible.

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, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

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 MINECO under grants AYA2015-71825, ESP2015-88861, FPA2015-68048, SEV-2012-0234, 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 Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020.

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. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish



or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities

for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. This research manuscript made use of Astropy [Astropy Collaboration \(2013, 2018\)](#), [GetDist Lewis \(2019\)](#) and [Matplotlib Hunter \(2007\)](#), and has been prepared using NASA's Astrophysics Data System Bibliographic Services.

- 
- Aihara H., et al. 2018, PASJ, 70, S4  
 Alam S., et al. 2015, The Astrophysical Journal Supplement Series, 219, 12  
 Alam S., et al. 2021, Phys. Rev. D, 103, 083533  
 Alarcon A., et al. 2020, MNRAS, 498, 2614  
 Amon A., et al., 2021, To be submitted to PRD  
 Asgari M., et al. 2017, MNRAS, 464, 1676  
 Asgari M., et al. 2021, A&A, 645, A104  
 Astropy Collaboration 2013, A&A, 558, A33  
 Astropy Collaboration 2018, A.J., 156, 123  
 Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625  
 Bartelmann M., Schneider P., 2001, Phys. Rept., 340, 291  
 Becker M. R., 2013, MNRAS, 435, 115  
 Benjamin J., et al. 2007, MNRAS, 381, 702  
 Bernstein G. M., et al. 2016, MNRAS, 459, 4467  
 Bernstein G. M., et al. 2018, PASP, 130, 054501  
 Bertin E., Arnouts S., 1996, AAPS, 117, 393  
 Blazek J., Vlah Z., Seljak U., 2015, JCAP, 2015, 015  
 Blazek J. A., et al. 2019, Phys. Rev. D, 100, 103506  
 Bridle S., King L., 2007, New Journal of Physics, 9, 444  
 Bridle S., et al. 2010, MNRAS, 405, 2044  
 Brown M. L., et al. 2002, MNRAS, 333, 501  
 Brown M. L., et al. 2003, MNRAS, 341, 100  
 Buchs R., Davis C., et al., 2019, MNRAS, 489, 820  
 Capozzi F., et al. 2014, Phys. Rev. D, 89, 093018  
 Carretero J., et al. 2015, MNRAS, 447, 646  
 Catelan P., Porciani C., 2001, MNRAS, 323, 713  
 Catelan P., Kamionkowski M., Blandford R. D., 2001, MNRAS, 320, L7  
 Cawthon R., et al., 2020, Submitted to MNRAS  
 Chang C., et al. 2013, MNRAS, 434, 2121  
 Chang C., et al. 2019, MNRAS, 482, 3696  
 Cordero J. P., Harrison I., et al., 2021, arXiv e-prints, p. arXiv:2109.09636  
 Crittenden R. G., et al. 2002, Astrophys. J., 568, 20  
 Crocce M., et al. 2015, MNRAS, 453, 1513  
 van Daalen M. P., et al. 2011, MNRAS, 415, 3649  
 van Daalen M. P., et al. 2014, MNRAS, 440, 2997  
 Dark Energy Survey Collaboration 2016, Phys. Rev. D, 94, 022001  
 Dark Energy Survey Collaboration 2018, Phys. Rev. D, 98, 043526  
 Dark Energy Survey Collaboration 2019, Phys. Rev. D, 99, 123505  
 Dark Energy Survey Collaboration 2020, Phys. Rev. D, 102, 023509  
 Dark Energy Survey Collaboration 2021b, To be submitted to PRD  
 Dark Energy Survey Collaboration 2021a, To be submitted to PRD  
 De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, MNRAS, 459, 3078  
 DeRose J., et al. 2019a, arXiv e-prints, p. arXiv:1901.02401  
 DeRose J., et al. 2019b, Astrophys. J., 875, 69  
 DeRose J., et al., 2021a, To be submitted to MNRAS  
 DeRose J., Becker M. R., Wechsler R. H., 2021b, arXiv e-prints, p. arXiv:2105.12104  
 Dodelson S., Shapiro C., White M., 2006, Phys. Rev. D, 73, 023009  
 Drlica-Wagner A., et al. 2018, ApJS, 235, 33  
 Dubois Y., et al. 2014, MNRAS, 444, 1453  
 Dubois Y., et al. 2016, MNRAS, 463, 3948  
 Eckert K., et al. 2020, MNRAS, 497, 2529  
 Elvin-Poole J., MacCrann N., et al., 2021, To be submitted to MNRAS  
 Erben T., et al., 2013, MNRAS, 433, 2545  
 Esteban I., et al. 2019, Journal of High Energy Physics, 2019, 106  
 Euclid Collaboration 2019, MNRAS, 484, 5509  
 Everett S., et al., 2020, Submitted to ApJS  
 Faber S. M., et al. 2007, Astrophys. J., 665, 265  
 Fang X., et al. 2017, JCAP, 2017, 030  
 Fang X., Eifler T., Krause E., 2020, MNRAS, 497, 2699  
 Flaughner B., et al. 2015, A.J., 150, 150  
 Fortuna M. C., et al. 2021, MNRAS, 501, 2983  
 Fosalba P., et al. 2015a, MNRAS, 447, 1319  
 Fosalba P., et al. 2015b, MNRAS, 448, 2987  
 Freedman W. L., et al. 2019, Astrophys. J., 882, 34  
 Friedrich O., et al., 2020, Submitted to MNRAS  
 Frieman J. A., Turner M. S., Huterer D., 2008, ARAA, 46, 385  
 Fu L., et al., 2008, Astron. Astrophys., 479, 9  
 Gatti M., et al. 2018, MNRAS, 477, 1664  
 Gatti M., Giannini G., et al., 2020, Submitted to MNRAS  
 Gatti M., Sheldon E., et al., 2021, MNRAS, 504, 4312  
 Giblin B., et al. 2021, A&A, 645, A105  
 Hamana T., et al., 2003, Astrophys. J., 597, 98  
 Hamana T., et al. 2020, PASJ, 72, 16  
 Handley W., 2019, The Journal of Open Source Software, 4, 1414  
 Handley W., Lemos P., 2019a, Phys. Rev. D, 100, 023512  
 Handley W., Lemos P., 2019b, Phys. Rev., D100, 043504  
 Handley W. J., Hobson M. P., Lasenby A. N., 2015, MNRAS, 453, 4384  
 Harnois-Déraps J., et al. 2015, MNRAS, 450, 1212  
 Hartley W. G., Choi A., et al., 2020, Submitted to MNRAS  
 Hettterscheidt M., et al. 2007, Astron. Astrophys., 468, 859  
 Heymans C., et al., 2005, MNRAS, 361, 160  
 Heymans C., et al. 2006, MNRAS, 368, 1323  
 Heymans C., et al. 2012, MNRAS, 427, 146  
 Heymans C., et al. 2021, A&A, 646, A140  
 Hikage C., et al. 2019, PASJ, 71, 43  
 Hildebrandt H., et al. 2020, A&A, 633, A69  
 Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526  
 Hirata C. M., et al. 2007, MNRAS, 381, 1197  
 Hoekstra H., et al. 2002, Astrophys. J., 572, 55  
 Hoffmann K., et al. 2015, MNRAS, 447, 1724  
 Hoffmann K., et al. 2021, *in preparation*  
 Hu W., 1999, The Astrophysical Journal, 522, L21–L24  
 Hu W., Jain B., 2004, Phys. Rev. D, 70, 043009  
 Huang H.-J., et al. 2021, MNRAS, 502, 6010  
 Huff E., Mandelbaum R., 2017, arXiv e-prints, p. arXiv:1702.02600  
 Huff E. M., et al. 2014a, MNRAS, 440, 1296  
 Huff E. M., et al. 2014b, MNRAS, 440, 1322  
 Hunter J. D., 2007, Computing in Science & Engineering, 9, 90  
 Huterer D., 2002, Phys. Rev. D, 65, 063001  
 Huterer D., et al. 2006, MNRAS, 366, 101

- Jain B., Seljak U., 1997, *Astrophys. J.*, 484, 560
- Jarvis M., et al. 2003, *Astron. J.*, 125, 1014
- Jarvis M., Bernstein G., Jain B., 2004, *MNRAS*, 352, 338
- Jarvis M., et al. 2006, *Astrophys. J.*, 644, 71
- Jarvis M., et al., 2021, *MNRAS*, 501, 1282
- Jee M. J., et al. 2013, *Astrophys. J.*, 765, 74
- Jee M. J., et al. 2016, *Astrophys. J.*, 824, 77
- Jeffreys H., 1935, *Proceedings of the Cambridge Philosophical Society*, 31, 203
- Joachimi B., et al. 2011, *A&A*, 527, A26
- Joachimi B., et al. 2013a, *MNRAS*, 431, 477
- Joachimi B., et al. 2013b, *MNRAS*, 436, 819
- de Jong J. T. A., et al. 2013, *The Messenger*, 154, 44
- Kaiser N., Wilson G., Luppino G. A., 2000, arXiv:astro-ph/0003338,
- Khandai N., et al. 2015, *MNRAS*, 450, 1349
- Kilbinger M., et al., 2013, *MNRAS*, 430, 2200
- Kitching T. D., et al. 2012, *MNRAS*, 423, 3163
- Kitching T. D., et al., 2014, *MNRAS*, 442, 1326
- Krause E., Eifler T., 2017, *MNRAS*, 470, 2100
- Krause E., Eifler T., Blazek J., 2016, *MNRAS*, 456, 207
- Krause E., et al., 2021, To be submitted to PRD
- Kullback S., Leibler R. A., 1951, *Ann. Math. Statist.*, 22, 79
- Laigle C., et al. 2016, *ApJS*, 224, 24
- Lawrence E., et al. 2017, *Astrophys. J.*, 847, 50
- Lazanu A., et al. 2016, *Phys. Rev. D*, 93, 083517
- Leauthaud A., et al., 2007, *Astrophys. J. Suppl.*, 172, 219
- Lemos P. Raveri M., et al., 2020, Submitted to MNRAS
- Lewis A., 2019, arXiv e-prints, p. arXiv:1910.13970
- Lewis A., Challinor A., Lasenby A., 2000, *Astrophys. J.*, 538, 473
- Limber D. N., 1953, *Astrophys. J.*, 117, 134
- Lin H., et al. 2012, *Astrophys. J.*, 761, 15
- LoVerde M., Afshordi N., 2008, *Phys. Rev. D*, 78, 123506
- MacCrann N., et al., 2020, Submitted to MNRAS
- Mandelbaum R., et al. 2006, *MNRAS*, 367, 611
- Mandelbaum R., et al. 2014, *ApJS*, 212, 5
- Mandelbaum R., et al. 2018a, *PASJ*, 70, S25
- Mandelbaum R., et al. 2018b, *MNRAS*, 481, 3170
- Mantz A. B., et al. 2015, *MNRAS*, 446, 2205
- Martinelli M., et al. 2021, *A&A*, 649, A100
- Massey R., et al., 2007a, *Astrophys. J. Suppl.*, 172, 239
- Massey R., et al. 2007b, *MNRAS*, 376, 13
- McEwen J. E., et al. 2016, *JCAP*, 2016, 015
- Mead A., et al. 2015, *MNRAS*, 454, 1958
- Miller L., et al., 2013, *MNRAS*, 429, 2858
- Miralda-Escude J., 1991, *Astrophys. J.*, 380, 1
- Morganson E., et al. 2018, *PASP*, 130, 074501
- Muir J., et al. 2020, *MNRAS*,
- Myles J., Alarcon A., et al., 2020, Submitted to MNRAS
- Okumura T., Jing Y. P., Li C., 2009, *Astrophys. J.*, 694, 214
- Pandey S., et al., 2021, To be submitted to MNRAS
- Planck Collaboration 2020a, *A&A*, 641, A5
- Planck Collaboration 2020b, *A&A*, 641, A6
- Planck Collaboration 2020c, *A&A*, 641, A8
- Porredon A., et al., 2021a, To be submitted to PRD
- Porredon A., et al., 2021b, *Phys. Rev. D*, 103, 043503
- Prat J., et al. 2019, *MNRAS*, 487, 1363
- Prat J., et al., 2021, To be submitted to PRD
- Raveri M., Hu W., 2019a, *Phys. Rev. D*, 99, 043506
- Raveri M., Hu W., 2019b, *Phys. Rev. D*, 99, 043506
- Raveri M., Zacharegkas G., Hu W., 2020, *Phys. Rev. D*, 101, 103527
- Refregier A., Rhodes J., Groth E. J., 2002, *Astrophys. J. Lett.*, 572, L131
- Rhodes J., et al. 2004, *Astrophys. J.*, 605, 29
- Riess A. G., et al. 2021, *Astrophys. J. Lett.*, 908, L6
- Rozo E., et al. 2016, *MNRAS*, 461, 1431
- Samuroff S., et al. 2017, *MNRAS*, 465, L20
- Samuroff S., et al. 2019, *MNRAS*, 489, 5453
- Sánchez C., Bernstein G. M., 2019, *MNRAS*, 483, 2801
- Sánchez C., Prat J., et al., 2021, To be submitted to PRD
- Scarlata C., et al. 2007, *ApJS*, 172, 406
- Schaye J., et al. 2010, *MNRAS*, 402, 1536
- Schaye J., et al. 2014, *MNRAS*, 446, 521
- Schaye J., et al. 2015, *MNRAS*, 446, 521
- Schmidt F., et al. 2009, *Astrophys. J.*, 702, 593
- Schmitz D. M., et al. 2018, *JCAP*, 2018, 030
- Schneider P., Lombardi M., 2003, *A&A*, 397, 809
- Schneider P., van Waerbeke L., Mellier Y., 2002, *A&A*, 389, 729
- Schneider P., Eifler T., Krause E., 2010, *A&A*, 520, A116
- Schneider A., et al. 2019, *Journal of Cosmology and Astroparticle Physics*, 2019, 020
- Schrabback T., et al., 2010, *Astron. Astrophys.*, 516, A63
- Scolnic D. M., et al. 2018, *Astrophys. J.*, 859, 101
- Seitz C., Schneider P., 1997, *A&A*, 318, 687
- Semboloni E., et al. 2011, *MNRAS*, 417, 2020
- Semboloni E., Hoekstra H., Schaye J., 2013, *MNRAS*, 434, 148
- Sevilla-Noarbe I., et al., 2020, Submitted to ApJS
- Shapiro C., 2009, *Astrophys. J.*, 696, 775
- Sheldon E. S., Huff E. M., 2017, *Astrophys. J.*, 841, 24
- Singh S., Mandelbaum R., 2016, *MNRAS*, 457, 2301
- Singh S., Mandelbaum R., More S., 2015, *MNRAS*, 450, 2195
- Smith R. E., et al. 2003, *MNRAS*, 341, 1311
- Springel V., 2005, *MNRAS*, 364, 1105
- Takahashi R., et al. 2012, *Astrophys. J.*, 761, 152
- Takahashi R., et al. 2020, *Astrophys. J.*, 895, 113
- Tegmark M., Zaldarriaga M., 2002, *Phys. Rev. D*, 66, 103508
- Tröster T., et al. 2020, *A&A*, 633, L10
- Troxel M. A., et al. 2018, *Phys. Rev. D*, 98, 043528
- Tugendhat T. M., Schäfer B. M., 2018, *MNRAS*, 476, 3460
- Van Waerbeke L., et al. 2000, *A&A*, 358, 30
- Vlah Z., Chisari N. E., Schmidt F., 2020, *JCAP*, 2020, 025
- Vogelsberger M., et al. 2014, *MNRAS*, 444, 1518
- Wechsler R. H., et al. 2021, arXiv e-prints, p. arXiv:2105.12105
- Wittman D. M., et al. 2000, *Nature (London)*, 405, 143
- Wright A. H., et al. 2020, *A&A*, 640, L14
- Yoon M., Jee M. J., 2021, *Astrophys. J.*, 908, 13
- Yoon M., et al. 2019, *Astrophys. J.*, 870, 111
- Zuntz J., et al. 2015, *Astronomy and Computing*, 12, 45
- Zuntz J., et al. 2018, *MNRAS*, 481, 1149

#### Appendix A: Construction of the MICE source catalog and results

In this section, we discuss the construction of a mock weak lensing source catalog with the MICE N-body simulation. Significantly, in addition to cosmological shear, these mock catalogs contain a realistic IA component, making them a good testing ground for our modeling; this is the primary motivation for using them here, as discussed in Sec. VB. The implementation of the IA signal into MICE is described in far more detail in Hoffmann et al. (2021), and is summarized below.

Each MICE galaxy has a (projected) shape assigned to it:

$$\gamma = \gamma_G + \gamma_I + \epsilon. \quad (\text{A1})$$

The cosmological part of this is obtained via ray tracing (see Fosalba et al. 2015b);  $\gamma_I$  is the coherent intrinsic component,

which is generated via semi-analytic modeling (see below and Hoffmann et al. 2021), and  $\epsilon$  is random Gaussian shape noise.

For the purposes of our test, it is desirable that the overall constraining power of the mock catalog (in terms of the figure-of-merit in the  $S_8 \times \Omega_m$  plane) resembles that of the fiducial source catalog; to that end, we match the constraining power of Y3 cosmic shear by adding a shape noise component to the covariance matrix of the hypothetical MICE survey, generated with COSMOLIKE. This scatter is chosen such that the uncertainty in the shape noise dominated regime is equivalent between MICE and METACAL:  $\Delta C_{\text{METACAL}}^{\kappa\kappa}(\ell) = \Delta C_{\text{MICE}}^{\kappa\kappa}(\ell)$ , with

$$\Delta C^{\kappa\kappa}(\ell) \sim \frac{1}{\sqrt{f_{\text{sky}}}} \left( \frac{\sigma_e^2}{\bar{n}} \right) \quad (\text{A2})$$

where  $f_{\text{sky}}$  is the observed fraction of the total sky area,  $\bar{n}$  is the (effective) galaxy number density and  $\sigma_e^2$  is the variance of the ellipticities. With  $\bar{n}$  and  $f_{\text{sky}}$  fixed in both catalogs (our MICE catalog has approximately 130M galaxies over an area of 5191 deg<sup>2</sup> while METACAL has 100M galaxies spanning an area of 4143 deg<sup>2</sup>), and  $\sigma_e = 0.261$  fixed in METACAL, we adjust  $\sigma_{e,\text{MICE}}$  so that the equality is reached, and verify that we achieve a comparable constraining power as DES Y3 in the 2D  $S_8 \times \Omega_m$ . The IA contribution to the MICE shears ( $\gamma_I$ ) has, additionally, an irreducible noise coming from a randomization of galaxy orientations. This randomization is a component of the IA modeling and has been tuned to match observed IA correlations from the BOSS LOWZ sample (Singh & Mandelbaum 2016).

It is also desirable that the source galaxy photometric properties in the mock also resemble those of galaxies in the fiducial DES Y3 catalog, since IAs can depend on galaxy color and luminosity. In summary, we verify that the distribution of *riz* magnitudes and color-magnitude distributions are similar between MICE and METACALIBRATION. It would also be beneficial if the redshift distributions  $n(z)$  on the mock catalog were closely matched to what is found in real data (Fig. 3). Since the implementation in MICE of the SOMPZ framework utilized in DES Y3 (Myles, Alarcon et al. 2020; Buchs, Davis et al. 2019) is beyond the scope of this work, and since the goal of this model test should be agnostic to the exact shape of the redshift distributions, we utilize Directional Neighborhood Fitting (DNF) photo-*z*s (De Vicente et al. 2016) for our mock  $n(z)$ s, as these are made available by default with the MICE data releases. In general, we find that DNF redshift distributions are narrower and less overlapping across bins than the fiducial DES Y3  $n(z)$ s.

With the mock catalog constructed, we then apply the same pipelines that were used for measuring the tomographic cosmic shear data vector  $\xi_{\pm}^{ij}(\theta)$  and inferring cosmological parameters. Where possible we mirror the model choices and nuisance parameter marginalization used in the analysis of the real data.

We measure 2 sets of shear correlation functions from the mock: a ‘‘MICE Baseline’’ data vector estimated from the IA-free GG signal (i.e. using only the  $\gamma^G$  component of Eq. A1), and a ‘‘MICE IA’’ data vector which includes the full GG+GI+II signal, as well shape noise added at the catalog

level. Since the former captures the cosmological signal (plus cosmic variance), it provides a useful fiducial case, relative to which we can gauge biases. The input cosmology is known for MICE, however the IA signal is, by construction, not exactly mapped into a set of NLA or TATT parameters. Thus, for a single realization of the mock, our expectations are: 1) that the constraints on the MICE Baseline GG-only vector are consistent with the input cosmology while at the same time the IA parameter constraints are consistent with zero; 2) that cosmology constraints on the full MICE IA data vector are not excessively biased with respect to the baseline constraints.

Both our expectations are fulfilled, with MICE IA offset by  $\sim 0.6\sigma$  from the peak of Baseline. Our results are shown in Fig. 14. Although this shift is larger than the  $0.3\sigma$  tolerance we applied to previous tests on analytic data, it is worth bearing in mind that the nature of the test discussed here is slightly different. That is, we only have one realization of the noise in the IA vector (both cosmic variance and shape noise). While this test is valuable, in the sense that large shifts would be a sign of significant problems in the fidelity of our model, offsets within  $\sim 1 - 2\sigma$  are entirely consistent with noise.

We also run the inference pipeline on the full IA data vector while fixing the cosmological parameters at the true MICE input to isolate the IA signal. Those are the ‘‘reduced’’ parameter space curves in Fig. 14. While constraints obtained this way are artificially tight, the exercise shows that the IA parameters measured in the reduced case are consistent within  $1\sigma$  with parameters obtained while varying cosmology. Additionally, we do find a marginal preference for a negative  $a_1$  parameter in both cases that involve nonzero IA parameters. Although this is likely not a physical effect, and nor in this case can it be the result of unmodeled photo-*z* error, it is quite conceivable that it is simply a result of a small but positive ‘‘true’’ IA signal, combined with shape noise.

This test on MICE is aimed at verifying that our fiducial model can recover unbiased cosmological constraints from a data vector that includes intrinsic alignments and that is not generated by an analytic pipeline. While we do obtain reasonable results and the present test has the rare benefit of including a realistic, simulated ( $N$ -body with semi-analytic model) IA signal, it is statistically weaker than a more complete validation of DES pipelines on simulations (DeRose et al. 2021a), mainly due to the lack of independent realizations. A complete analysis on a relatively large number of MICE mocks is left for future work.

## Appendix B: Robustness of IA to Unmodeled Redshift Error

In this Appendix we set out in more detail the tests on analytic data introduced in Sec. VA3. These tests, which are based on noiseless analytic data, are designed to test the specific question of whether our choice of IA model behaves well in the presence of realistic errors in the redshift distributions. We focus here on synthetic data vectors which do not contain real data complexity, in which case a thorough testing of redshift calibration is more complicated and performed by our companion Amon et al. (2021). We use the ensemble of 6000

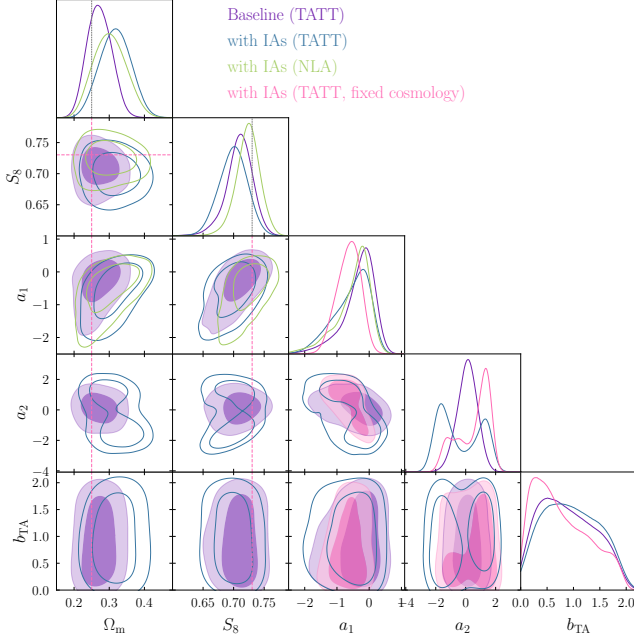


FIG. 14. MICE posteriors: The baseline case (purple) includes only the GG part of the signal (no IAs) and IA parameters are consistent with zero. The contours labelled “with IAs” (blue, green, pink) includes GG+GI+II, with the IA model indicated in parentheses. Note that two IA redshift power law indices are included in the fits, as per the fiducial Y3 model; they are not shown here as they are only very weakly constrained, and the contours are uninformative. Dashed pink lines are the input MICE cosmology.

SOMPZ realizations of  $n(z)$ 's, described in [Myles, Alarcon et al. \(2020\)](#); for each sample of the Y3 redshift distributions, we compute a simulated cosmic shear data vector  $\xi_{\pm}^{\text{sample}}(\theta)$  at a fixed set of input cosmological parameters. Using the data covariance, we can then calculate the  $\Delta\chi^2$  between  $\xi_{\pm}^{\text{sample}}(\theta)$  and a similarly generated cosmic shear data vector, but with the fiducial (mean)  $n(z)$  as input,  $\xi_{\pm}^{\text{mean}}(\theta)$ . We find that the distribution of  $\Delta\chi^2$  obtained for the 6000 samples peaks close to zero ( $\Delta\chi^2 < 1.0$ , corresponding to small perturbations of the data vector) and has a long tail out to more extreme cases ( $\Delta\chi^2 \sim 50.0$ ). Given our finite computing resources, we do not run nested sampling chains on all of the 6000 scenarios, but rather choose three of increasing severity, corresponding to the 60th, 95th and 99th percentiles of the  $p(\Delta\chi^2)$  distribution.

We next translate these  $\Delta\chi^2$  errors into cosmological biases by analysing the three  $\xi_{\pm}^{\text{sample}}(\theta)$  vectors with our fiducial setup. Although that setup includes redshift nuisance parameters  $\Delta z^i$ , it cannot explicitly account for perturbations in the  $n(z)$  shape that don't map well into the first moment of the  $n(z)$  distribution. More extensive testing of this is presented in [Amon et al. \(2021\)](#); [Cordero, Harrison et al. \(2021\)](#).

The results of this exercise are shown in Fig. 15. Each of the unfilled contours represents a different scenario of redshift error, and can be compared with the solid purple (which has no redshift error). In this limited testing scenario with synthetic data we see no large biases in the IA model parameters or

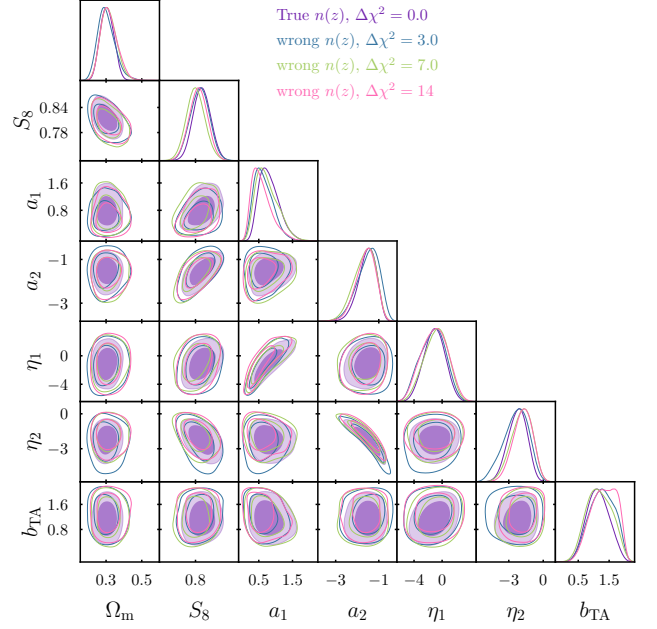


FIG. 15. Analysis of synthetic data generated with the inclusion of unmodeled redshift errors. The purple contour shows the baseline posterior, obtained by analysing a simulated TATT data vector with the correct  $n(z)$ ; the three other colours show redshift error scenarios of varying extremity. In each case the difference in  $\chi^2$  induced by the redshift error, at the input point in parameter space, is indicated in the legend.

$S_8$ , which are recovered correctly, and no hints that the TATT model parameters are absorbing the unmodeled error.

### Appendix C: Comparing NLA and TATT in DES Y3

In this appendix we explore in more detail the differences between our fiducial IA model for Y3 (TATT), and the more commonly used NLA model. As discussed in Sec. IV D 2, TATT is a physically-motivated model containing NLA, such that in some limit ( $a_2, b_{\text{TA}} \rightarrow 0$ ), the two are the same. Amongst various other permutations, we fit both the two-parameter NLA model and the five-parameter TATT model to the Y3 cosmic shear data. Using the more complex IA model results in a slight widening of the contours, as can be seen in Fig. 8, and also a rotation in the direction of the  $\sigma_8 - \Omega_m$  degeneracy. These two effects combine to produce a  $\sim 17\%$  difference in the projected error bar on  $S_8$  between the two models. When considering the 2D Figure of Merit in the  $S_8 - \Omega_m$  plane, the difference is  $\sim 25\%$ .

In addition to the difference in constraining power, NLA and TATT result in qualitatively different predictions for the IA signal. We illustrate this in Fig. 16, which shows the theory IA contributions, generated using the respective best fitting IA parameters from the two chains. Both the absolute signal, and the contribution relative to the cosmological shear (GG) are included in the two panels. Although the choice of model

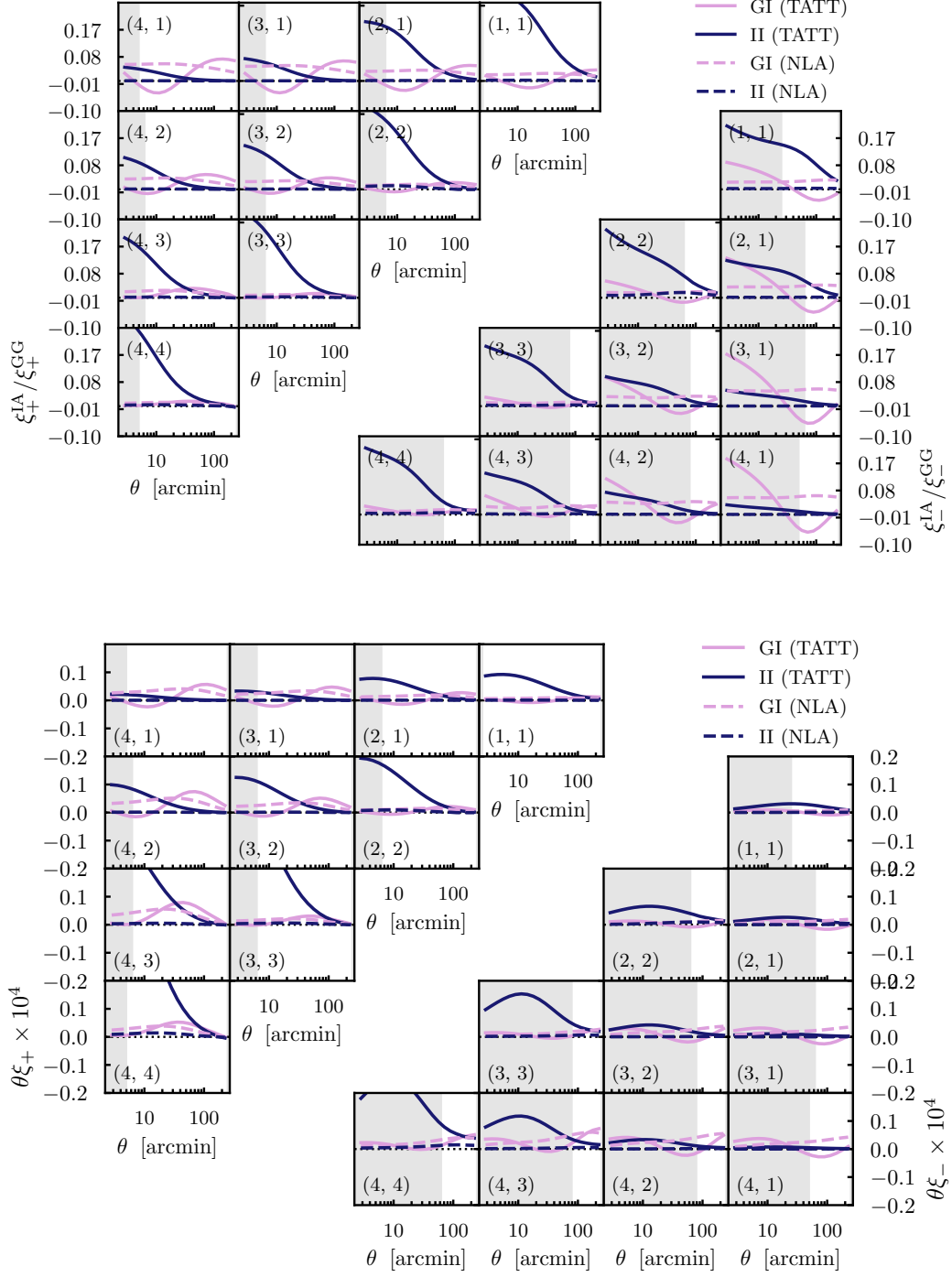


FIG. 16. The best fitting IA predictions, based on NLA and TATT analyses. In the two panels we show the fractional IA contributions relative to the cosmological signal (top), and the absolute contributions (bottom). Both IA components, GI and II, are plotted separately (blue and light purple). As before, shaded regions indicate scales discarded from the analyses.

does not change the conclusion that IAs are subdominant in all bins (at the level of a few percent), the shape and sign of the IA correlations do differ somewhat. Strikingly, TATT predicts a much stronger II signal than NLA, which is largely driven by the tidal torque ( $a_2$ ) part of the model. It is also worth remarking here that although they look different, neither of these IA scenarios has been convincingly ruled out by direct measurements. That is, if one generates projected intrinsic alignment correlations (see e.g. [Mandelbaum et al. 2006](#) Sec. 3.2 for the full definition)  $w_{g+}$  and  $w_{++}$  predictions, using the IA power spectra from these fits, the results are within the range allowed by measurements on faint blue and red galaxies at low redshift from SDSS ([Mandelbaum et al. 2006](#), [Singh et al. 2015](#)).

#### Appendix D: Full Posterior Constraints

We show a set of 2D projections of our cosmological and IA parameter posteriors in Fig. 17, along with their priors. We note that  $\sigma_8$  and  $S_8$  are derived parameters which we do not sample over, so projections make their apparent priors non-uniform. Many of the parameters in our analysis are either prior dominated, or not of physical interest, and so are not shown in the main part of the paper. They are included here

for completeness.

#### Appendix E: Changes After Unblinding

Three changes to our analysis happened after unblinding, neither of them affecting significantly the DES Y3 cosmic shear cosmological constraints, particularly  $S_8$ . The first is a planned modification to the covariance matrix: after the tests described in Sec. V succeed and cosmology constraints are obtained, the analytic covariance matrix is re-computed at the best-fit (maximum posterior) parameters of the full  $3 \times 2$ pt data, and our cosmological constraints are updated with new nested sampling chains.

The second change involves the DES Y3 lens sample, which only enters the cosmic shear data via lensing ratios (Sec. IV F). We substitute the original REDMAGIC sample by the magnitude-limited MAGLIM sample. We point the reader to [Dark Energy Survey Collaboration \(2021a\)](#) for further details and motivation for this change.

The third change is decision for presenting the  $\Lambda$ CDM-Optimized scale cuts in cosmic shear as one of the main DES Y3 results. While fundamentally the optimized scale cuts also lead to robust  $< 0.3\sigma$  shifts in our simulated analysis and therefore could have been a reasonable pre-unblinding choice on a  $\xi_{\pm}$ -only analysis, this was only decided after unblinding. We report that here for transparency.

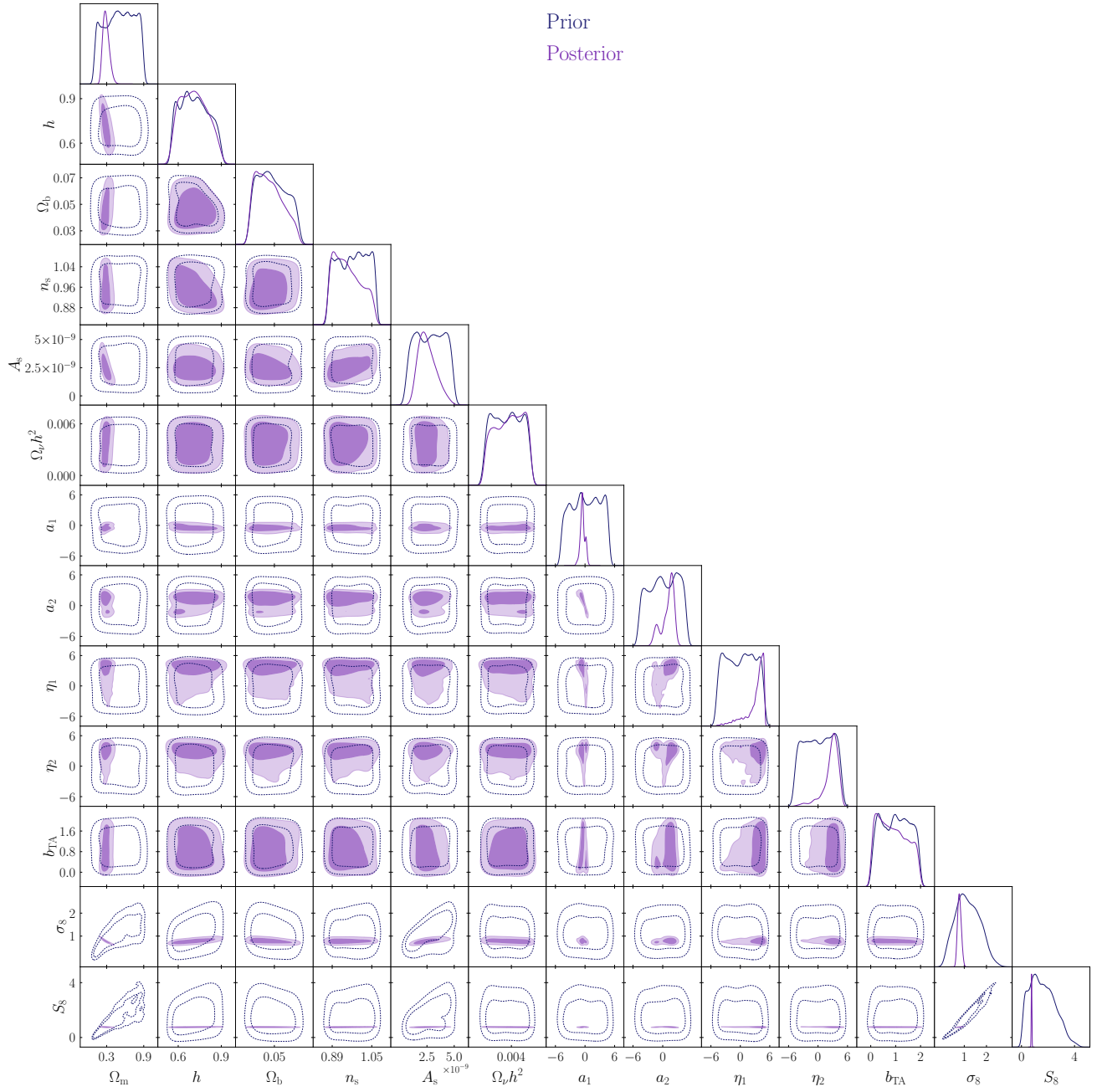


FIG. 17. Posterior distributions (purple solid) and priors (blue dashed) of our analysis. We intentionally leave out hard-boundary corrections on the distributions above for ease of visualization. Our 1D marginalized priors are flat in the cosmological parameters described in Table I, but sampling and 1-dimensional marginalization can make them look noisy and non-uniform.