

# Dark Energy Survey Year 1 Results: Cosmological constraints from cluster abundances and weak lensing

T. M. C. Abbott,<sup>1</sup> M. Aguena,<sup>2,3</sup> A. Alarcon,<sup>4,5</sup> S. Allam,<sup>6</sup> S. Allen,<sup>7</sup> J. Annis,<sup>6</sup> S. Avila,<sup>8</sup> D. Bacon,<sup>9</sup> K. Bechtol,<sup>10,11</sup> A. Bermeo,<sup>12</sup> G. M. Bernstein,<sup>13</sup> E. Bertin,<sup>14,15</sup> S. Bhargava,<sup>12</sup> S. Bocquet,<sup>16,17</sup> D. Brooks,<sup>18</sup> D. Brout,<sup>13</sup> E. Buckley-Geer,<sup>6</sup> D. L. Burke,<sup>19,20</sup> A. Carnero Rosell,<sup>21</sup> M. Carrasco Kind,<sup>22,23</sup> J. Carretero,<sup>24</sup> F. J. Castander,<sup>4,5</sup> R. Cawthon,<sup>11</sup> C. Chang,<sup>25,26</sup> X. Chen,<sup>27</sup> A. Choi,<sup>28</sup> M. Costanzi,<sup>29,30</sup> M. Crocce,<sup>4,5</sup> L. N. da Costa,<sup>3,31</sup> T. M. Davis,<sup>32</sup> J. De Vicente,<sup>21</sup> J. DeRose,<sup>7,19</sup> S. Desai,<sup>33</sup> H. T. Diehl,<sup>6</sup> J. P. Dietrich,<sup>17</sup> S. Dodelson,<sup>34</sup> P. Doel,<sup>18</sup> A. Drlica-Wagner,<sup>25,6,26</sup> K. Eckert,<sup>13</sup> T. F. Eifler,<sup>35,36</sup> J. Elvin-Poole,<sup>28,37</sup> J. Estrada,<sup>6</sup> S. Everett,<sup>38</sup> A. E. Evrard,<sup>39,27</sup> A. Farahi,<sup>27</sup> I. Ferrero,<sup>40</sup> B. Flaugher,<sup>6</sup> P. Fosalba,<sup>4,5</sup> J. Frieman,<sup>6,26</sup> J. García-Bellido,<sup>8</sup> M. Gatti,<sup>24</sup> E. Gaztanaga,<sup>4,5</sup> D. W. Gerdes,<sup>39,27</sup> T. Giannantonio,<sup>41,42</sup> P. Giles,<sup>43</sup> S. Grandis,<sup>17</sup> D. Gruen,<sup>7,44,45</sup> R. A. Gruendl,<sup>22,23</sup> J. Gschwend,<sup>3,31</sup> G. Gutierrez,<sup>6</sup> W. G. Hartley,<sup>46,47</sup> S. R. Hinton,<sup>32</sup> D. L. Hollowood,<sup>38</sup> K. Honscheid,<sup>28,37</sup> B. Hoyle,<sup>48,49</sup> D. Huterer,<sup>27</sup> D. J. James,<sup>50</sup> M. Jarvis,<sup>13</sup> T. Jeltema,<sup>38</sup> M. W. G. Johnson,<sup>23</sup> M. D. Johnson,<sup>23</sup> S. Kent,<sup>6,26</sup> E. Krause,<sup>35</sup> R. Kron,<sup>6,26</sup> K. Kuehn,<sup>51,52</sup> N. Kuropatkin,<sup>6</sup> O. Lahav,<sup>46</sup> T. S. Li,<sup>53,54</sup> C. Lidman,<sup>55</sup> M. Lima,<sup>2,3</sup> H. Lin,<sup>6</sup> N. MacCrann,<sup>28,37</sup> M. A. G. Maia,<sup>3,31</sup> A. Mantz,<sup>44</sup> J. L. Marshall,<sup>56</sup> P. Martini,<sup>28,57</sup> J. Mayers,<sup>43</sup> P. Melchior,<sup>53</sup> J. Mena-Fernández,<sup>21</sup> F. Menanteau,<sup>22,23</sup> R. Miquel,<sup>58,24</sup> J. J. Mohr,<sup>17,48</sup> R. C. Nichol,<sup>59</sup> B. Nord,<sup>6</sup> R. L. C. Ogando,<sup>3,31</sup> A. Palmese,<sup>6,26</sup> F. Paz-Chinchón,<sup>22,23</sup> A. A. Plazas,<sup>53</sup> J. Prat,<sup>24</sup> M. M. Rau,<sup>34</sup> A. K. Romer,<sup>43</sup> A. Roodman,<sup>44,45</sup> P. Rooney,<sup>43</sup> E. Rozo,<sup>60</sup> E. S. Rykoff,<sup>44,45</sup> M. Sako,<sup>13</sup> S. Samuroff,<sup>34</sup> C. Sánchez,<sup>13</sup> E. Sanchez,<sup>21</sup> A. Saro,<sup>29,61,30</sup> V. Scarpine,<sup>6</sup> M. Schubnell,<sup>27</sup> D. Scolnic,<sup>62</sup> S. Serrano,<sup>4,5</sup> I. Sevilla-Noarbe,<sup>21</sup> E. Sheldon,<sup>63</sup> J. Allyn. Smith,<sup>64</sup> M. Smith,<sup>65</sup> E. Suchyta,<sup>66</sup> M. E. C. Swanson,<sup>23</sup> G. Tarle,<sup>27</sup> D. Thomas,<sup>59</sup> C. To,<sup>7</sup> M. A. Troxel,<sup>62</sup> D. L. Tucker,<sup>6</sup> T. N. Varga,<sup>48,49</sup> A. von der Linden,<sup>67</sup> A. R. Walker,<sup>1</sup> R. H. Wechsler,<sup>7,44,45</sup> J. Weller,<sup>48,49</sup> R. D. Wilkinson,<sup>43</sup> H. Wu,<sup>37</sup> B. Yanny,<sup>6</sup> Y. Zhang,<sup>6</sup> Z. Zhang,<sup>26</sup> and J. Zuntz<sup>68</sup>

(DES Collaboration)\*

<sup>1</sup>*Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile*

<sup>2</sup>*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP, 05314-970, Brazil*

<sup>3</sup>*Laboratório Interinstitucional de e-Astronomia—LINEA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ—20921-400, Brazil*

<sup>4</sup>*Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain*

<sup>5</sup>*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*

<sup>6</sup>*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, Illinois 60510, USA*

<sup>7</sup>*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, California 94305, USA*

<sup>8</sup>*Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

<sup>9</sup>*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, United Kingdom*

<sup>10</sup>*LSST, 933 North Cherry Avenue, Tucson, Arizona 85721, USA*

<sup>11</sup>*Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, Wisconsin 53706-1390, USA*

<sup>12</sup>*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, United Kingdom*

<sup>13</sup>*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

<sup>14</sup>*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*

<sup>15</sup>*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*

<sup>16</sup>*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, Illinois 60439, USA*

<sup>17</sup>*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany*

<sup>18</sup>*Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, United Kingdom*

<sup>19</sup>*Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, California 94305, USA*

<sup>20</sup>*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*

- <sup>21</sup>*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid 28040, Spain*
- <sup>22</sup>*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, Illinois 61801, USA*
- <sup>23</sup>*National Center for Supercomputing Applications, 1205 West Clark Street, Urbana, Illinois 61801, USA*
- <sup>24</sup>*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain*
- <sup>25</sup>*Department of Astronomy and Astrophysics, University of Chicago, Chicago, Illinois 60637, USA*
- <sup>26</sup>*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*
- <sup>27</sup>*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*
- <sup>28</sup>*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, Ohio 43210, USA*
- <sup>29</sup>*INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
- <sup>30</sup>*Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy*
- <sup>31</sup>*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ—20921-400, Brazil*
- <sup>32</sup>*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
- <sup>33</sup>*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
- <sup>34</sup>*Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15312, USA*
- <sup>35</sup>*Department of Astronomy/Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, Arizona 85721-0065, USA*
- <sup>36</sup>*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, California 91109, USA*
- <sup>37</sup>*Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA*
- <sup>38</sup>*Santa Cruz Institute for Particle Physics, Santa Cruz, California 95064, USA*
- <sup>39</sup>*Department of Astronomy, University of Michigan, Ann Arbor, Michigan 48109, USA*
- <sup>40</sup>*Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway*
- <sup>41</sup>*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom*
- <sup>42</sup>*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom*
- <sup>43</sup>*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, United Kingdom*
- <sup>44</sup>*Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, California 94305, USA*
- <sup>45</sup>*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*
- <sup>46</sup>*Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, United Kingdom*
- <sup>47</sup>*Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland*
- <sup>48</sup>*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany*
- <sup>49</sup>*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany*
- <sup>50</sup>*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, Massachusetts 02138, USA*
- <sup>51</sup>*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
- <sup>52</sup>*Lowell Observatory, 1400 Mars Hill Road, Flagstaff, Arizona 86001, USA*
- <sup>53</sup>*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, New Jersey 08544, USA*
- <sup>54</sup>*Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, California 91101, USA*
- <sup>55</sup>*The Research School of Astronomy and Astrophysics, Australian National University, ACT 2601, Australia*
- <sup>56</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, Texas 77843, USA*
- <sup>57</sup>*Department of Astronomy, The Ohio State University, Columbus, Ohio 43210, USA*
- <sup>58</sup>*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
- <sup>59</sup>*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, United Kingdom*
- <sup>60</sup>*Department of Physics, University of Arizona, Tucson, Arizona 85721, USA*
- <sup>61</sup>*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*

<sup>62</sup>*Department of Physics, Duke University Durham, North Carolina 27708, USA*<sup>63</sup>*Brookhaven National Laboratory, Bldg 510, Upton, New York 11973, USA*<sup>64</sup>*Austin Peay State University, Dept. Physics, Engineering and Astronomy,  
P.O. Box 4608 Clarksville, Tennessee 37044, USA*<sup>65</sup>*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, United Kingdom*<sup>66</sup>*Computer Science and Mathematics Division, Oak Ridge National Laboratory,  
Oak Ridge, Tennessee 37831*<sup>67</sup>*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*<sup>68</sup>*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, United Kingdom*

(Received 25 February 2020; accepted 15 June 2020; published 7 July 2020)

We perform a joint analysis of the counts and weak lensing signal of redMaPPer clusters selected from the Dark Energy Survey (DES) Year 1 dataset. Our analysis uses the same shear and source photometric redshifts estimates as were used in the DES combined probes analysis. Our analysis results in surprisingly low values for  $S_8 = \sigma_8(\Omega_m/0.3)^{0.5} = 0.65 \pm 0.04$ , driven by a low matter density parameter,  $\Omega_m = 0.179^{+0.031}_{-0.038}$ , with  $\sigma_8 - \Omega_m$  posteriors in  $2.4\sigma$  tension with the DES Y1 3x2pt results, and in  $5.6\sigma$  with the *Planck* CMB analysis. These results include the impact of post-unblinding changes to the analysis, which did not improve the level of consistency with other data sets compared to the results obtained at the unblinding. The fact that multiple cosmological probes (supernovae, baryon acoustic oscillations, cosmic shear, galaxy clustering and CMB anisotropies), and other galaxy cluster analyses all favor significantly higher matter densities suggests the presence of systematic errors in the data or an incomplete modeling of the relevant physics. Cross checks with x-ray and microwave data, as well as independent constraints on the observable-mass relation from Sunyaev-Zeldovich selected clusters, suggest that the discrepancy resides in our modeling of the weak lensing signal rather than the cluster abundance. Repeating our analysis using a higher richness threshold ( $\lambda \geq 30$ ) significantly reduces the tension with other probes, and points to one or more richness-dependent effects not captured by our model.

DOI: [10.1103/PhysRevD.102.023509](https://doi.org/10.1103/PhysRevD.102.023509)

## I. INTRODUCTION

The flat  $\Lambda$ CDM model, despite its apparent simplicity—six parameters suffice to define it—has proven able to describe a wide variety of observations, from the low to the high redshift Universe. Despite its successes, however, the two dominant components of the Universe in this model—the cold dark matter (CDM) and the cosmological constant ( $\Lambda$ )—lack a fundamental theory to connect them with the rest of physics. Ongoing (e.g., the Dark Energy Survey (DES),<sup>1</sup> Hyper Suprime-Cam,<sup>2</sup> Kilo-Degree Survey<sup>3</sup> eRosita,<sup>4</sup> South Pole Telescope (SPT),<sup>5</sup> Atacama Cosmology Telescope (ACT)<sup>6</sup>) and future surveys (e.g., Euclid,<sup>7</sup> Large Synoptic Survey Telescope,<sup>8</sup> WFIRST<sup>9</sup>) aim to further test the  $\Lambda$ CDM paradigm, as well as the mechanism that drives the cosmic

acceleration, be it a cosmological constant, some form of dark energy, or a modification of general relativity. Lacking a fundamental theory to test, one way to shed light on the latter is by looking at the evolution of cosmic structures over the past few Gyr, when the dark energy becomes dominant, and searching for discrepancies between the observables in the low-redshift Universe and the predictions for said observables derived from the high-redshift Universe as measured through observations of the cosmic microwave background (CMB) anisotropies [e.g., [1,2]].

The Dark Energy Survey is a six-year survey that mapped 5000 deg<sup>2</sup> of the southern sky in five broadband filters,  $g, r, i, z, Y$ , between August 2013 and January 2019, using the 570 megapixel Dark Energy Camera [DECam; [3]] mounted on the 4 m Blanco telescope at the Cerro Tololo Inter-American Observatory (CTIO). DES was designed with the primary goal of testing the  $\Lambda$ CDM model and studying the nature of dark energy through four key probes: cosmic shear, galaxy clustering, clusters of galaxies, and Type Ia supernovae.

Galaxy clusters have long proven to be a valuable cosmological tool: arising from the highest peaks of the matter density field, their abundance and spatial distribution are sensitive to the growth of structures and cosmic expansion [see e.g., [4,5], for reviews]. More specifically, the cluster abundance constrains the parameter combination

\*For comments or questions please contact: [des-publication-queries@listserv.fnal.gov](mailto:des-publication-queries@listserv.fnal.gov).

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

<sup>2</sup><http://hsc.mtk.nao.ac.jp/sspl/>.

<sup>3</sup><http://kids.strw.leidenuniv.nl/index.php>.

<sup>4</sup><http://www.mpe.mpg.de/eROSITA>.

<sup>5</sup><https://pole.uchicago.edu/>.

<sup>6</sup><https://act.princeton.edu/>.

<sup>7</sup><http://sci.esa.int/euclid/>.

<sup>8</sup><https://www.lsst.org/>.

<sup>9</sup><https://wfirst.gsfc.nasa.gov/>.

$\sigma_8(\Omega_m/0.3)^\alpha$ , where  $\Omega_m$  is the mean matter density of the Universe,  $\sigma_8$  is the present-day rms of the linear density field in spheres of  $8h^{-1}$  Mpc radius, and  $\alpha$  ranges between  $\sim 0.2$ – $0.5$  depending on the characteristics of the survey. The evolution of the cluster abundance can thus be used to measure the growth rate of cosmic structure, which in turn constrains dark energy and modified gravity models [e.g., [6–9]].

At present, cluster abundance studies at all wavelengths are limited by their ability to calibrate the relation between halo mass and the observable used as a mass proxy. Among the different techniques to calibrate the observable-mass relation, the weak lensing signal, based on the distortion of background galaxy images due to the gravitational lensing of intervening clusters, is the current gold standard [e.g., [7,9–11]]. Still, many sources of systematic uncertainty affect this type of measurement, including shear and photometric redshift biases, halo triaxiality, miscentering, and projection effects, each of which contribute a significant fraction of the total error budget [e.g., [12–15]]. As we will discuss later, this is especially true for the optically selected cluster sample adopted in this work, for which the systematic error represents  $\sim 60\%$  of the total error budget on mass estimates.

In this study we combine cluster abundances and weak-lensing mass estimates derived from data collected during the first year of observation of DES to simultaneously constrain cosmology and the observable-mass relation. Our optically selected catalog is built using the red sequence matched-filter probabilistic percolation cluster finder algorithm [redMaPPer; [16]]. For mass estimates, we rely on updated results of the stacked weak lensing analysis of [15], which include a new calibration of the selection effect bias.<sup>10</sup> The latter has been studied by means of numerical simulations by Wu *et al.* (in preparation) to validate the systematic bias correction adopted in [15]. The results of this analysis, which started before the unblinding but have been finalized only after, show that selection effects have a  $\sim 20$ – $30\%$  impact on stacked weak lensing mass measurements, a much larger effect compared to the  $\sim 4\%$  correction estimated in [15] combining simulations [17] and analytic estimates [18].

This analysis follows the methodology described in [19], in which we develop our pipeline using the redMaPPer SDSS cluster catalog. This analysis was performed blind to the cosmological parameters to avoid confirmation bias. However, the large tension between our original unblinded results and multiple cosmological probes, including *Planck* CMB [2], and especially the DES 3x2pt [20] results, motivated a careful review of our handling of systematics. This led us to revisit our estimates of the selection effects bias and, in turn, to reanalyze and update our results

post-unblinding. The analysis presented in the main text of the paper make use of this post-unblinding correction, and we will refer to it as the *unblinded* analysis. For completeness, the cosmological results obtained at the unblinding (*blinded* analysis, hereafter) are presented in the Appendix C. As discussed in the paper, the post-unblinding correction, while reducing by  $2\sigma$  the preferred  $\sigma_8$  value, does not improve the consistency of our posteriors with either the *Planck* CMB or the DES 3x2pt results.

This paper is organized as follows: In Sec. II we provide an overview of the DES Y1 data products used in this work. Section III presents the two data vectors—cluster abundance and mean weak-lensing mass estimates—employed for the cosmological analysis. Section IV describes our theoretical model to predict cluster counts and mean cluster masses, and thus derive cosmological and observable-mass relation parameter constraints. We present our results and address their consistency with other probes in Sec. V, while we discuss their implication in Sec. VI. Finally, we summarize and draw our conclusions in Sec. VII.

## II. DATA

In this work we use data collected by the DECam during the Year One (Y1) observational season, running from August 31, 2013 to February 9, 2014, which covers  $\sim 1800$  deg<sup>2</sup> of the southern sky in the *g*, *r*, *i*, *z* and *Y* bands [21]. Of the  $\sim 1800$  square degrees observed in Y1,  $\sim 17\%$  of them are excluded from the analysis due to a series of veto masks, vetting bright stars, bright nearby galaxies, globular clusters, and the Large Magellanic Cloud. The final DES Y1 footprint is shown in Fig. 1, and covers approximately 1500 deg<sup>2</sup> split in two non-contiguous regions: a larger region (1321 deg<sup>2</sup>; lower panel) overlapping the footprint of the South Pole Telescope Sunyaev-Zel’dovich Survey [22], and a smaller area (116 deg<sup>2</sup>; upper panel), which overlaps the Stripe-82 deep field of the Sloan Digital Sky Survey [SDSS, [23]].

In Secs. II A–II D we summarize the main data products used in this work, and refer the reader to the relevant papers for further details.

### A. The DES Y1 photometric catalog

Photometry and ‘clean’ galaxy samples are based on the Y1A1 gold catalog [21], the DES science-quality photometric catalog produced from Y1 data to enable cosmological analyses. This data set includes a multiband photometric object catalog as well as maps of survey depth, foreground masks, and star-galaxy classification. Galaxy fluxes are measured using the multiepoch, multi-object fitting (MOF) procedure described in [21]. The typical  $10\sigma$  limiting magnitude inside  $2''$  diameter apertures for galaxies in Y1A1 Gold using MOF photometry is  $g \simeq 23.7$ ,  $r \simeq 23.5$ ,  $i \simeq 22.9$ , and  $z \simeq 22.2$ . Due to its shallow depth and significant calibration uncertainty, the

<sup>10</sup>We use the term “selection effect bias” to refer to the bias introduced by the cluster finder for preferentially selecting clusters with properties that correlate with the lensing signal at fixed mass.



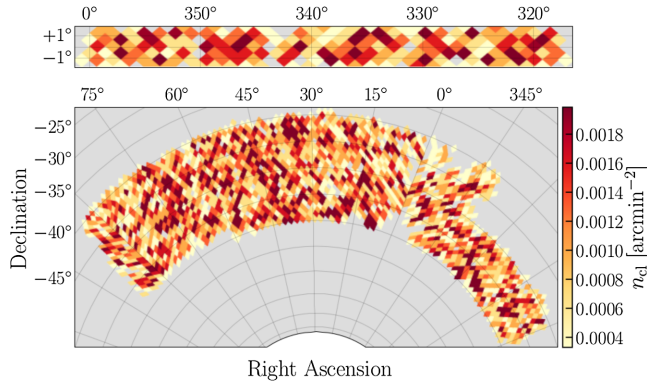


FIG. 1. The DES Y1 redMaPPer cluster density over the two non-contiguous regions of the Y1 footprint: the Stripe 82 region (116 deg<sup>2</sup>; upper panel) and the SPT region (1321 deg<sup>2</sup>; lower panel).

Y band photometry was used in neither the redMaPPer cluster finder nor for shape and photometric redshift measurements.

To build our cluster catalog, we rely on a subset of high-quality objects selected from the Y1A1 Gold catalog. First, we reject all objects classified as catalog artifacts, i.e., objects lying in regions having unphysical colors, astrometric discrepancies, or PSF model failures [Section 7.4 [21]]. The sample is further refined via the MODEST\_CLASS classifier, which was developed with the primary goal of selecting high-quality galaxy samples [Section 8.1 [21]]. Finally, only galaxies that are brighter in the  $z$  band than the local  $10\sigma$  limiting magnitude are included in the galaxy catalog used by the redMaPPer cluster finder.

## B. Cluster catalog and associated systematics

Our analysis relies on the DES Y1 redMaPPer cluster catalog. redMaPPer is a photometric cluster finding algorithm that identifies galaxy clusters as overdensities of red-sequence galaxies [16]. The algorithm has been extensively vetted against x-ray and Sunyaev-Zel'dovich (SZ) catalogs [24–27]. Incremental algorithmic updates are presented in [15,28,29]. Here, we present only a brief summary of the most salient features of the DES Y1 redMaPPer catalog. For further details on the algorithm, we refer the reader to the original work by [16].

The DES Y1 redMaPPer clusters are selected as overdensities of red-sequence galaxies in the DES Y1 photometric galaxy catalog. redMaPPer counts the excess number of red-sequence galaxies brighter than a specified luminosity threshold  $L_{\min}(z)$  within a circle of radius  $R_\lambda = 1.0h^{-1} \text{ Mpc}(\lambda/100)^{0.2}$ . This number of galaxies is called the *richness*, and is denoted as  $\lambda$ . We use all clusters of richness  $\lambda \geq 20$  in the present analysis. The catalog is locally volume limited in that we use the survey depth to determine the maximum redshift  $z_{\max}(\hat{n})$  at which galaxies

at our luminosity threshold are still detectable in the DES at  $10\sigma$ . Galaxy clusters are included in the volume-limited catalog if the cluster redshift  $z \leq z_{\max}(\hat{n})$ . The cluster survey footprint is mildly redshift dependent. It is defined as follows: a point  $\hat{n}$  in the sky at redshift  $z$  is included in the survey volume if a cluster at that redshift and position is masked by at most 20% by the galaxy mask. The above criteria, along with the recovered redshift distribution of the redMaPPer clusters, are used to generate a large random cluster catalog to characterize the survey volume.

A total of 7066 galaxy clusters are included in the DES Y1 redMaPPer volume-limited catalog. We remove 32 deg<sup>2</sup> corresponding to 10 noncontiguous deep fields for supernovae science, bringing down the total number of clusters to 6997. We further restrict ourselves to the redshift interval  $z \in [0.2, 0.65]$ , which reduces the number of galaxy clusters to 6504. redMaPPer performance below redshift  $z = 0.2$  is compromised by the lack of  $u$ -band data, while there are relatively few galaxy clusters in the catalog above redshift  $z = 0.65$ , making it a convenient upper limit for calculating binned abundances. Figure 1 shows the footprint of the DES Y1 redMaPPer cluster survey.

Galaxy clusters are centered on bright cluster galaxies, but not necessarily on the brightest cluster galaxy. The redMaPPer algorithm iteratively self-trains a filter that relies on galaxy brightness, cluster richness, and local galaxy density to determine candidate central galaxies. The algorithm centers the cluster on the most likely candidate central galaxy.

Turning to our characterization of systematic uncertainties in cluster finding, we note that, at a fundamental level, cluster catalogs should provide three measures for a cluster: (1) a sky location (center), (2) a cluster redshift estimate, and (3) an observable that serves as a proxy for mass. We briefly summarize the DES Y1 redMaPPer performance in each of these categories:

**Cluster centering:** The centering efficiency of the redMaPPer algorithm is studied using X-ray imaging by [30]. That work demonstrates that the fraction of correctly centered redMaPPer clusters is  $f_{\text{cen}} = 0.75 \pm 0.08$ . The distribution of radial offsets for miscentered clusters relative to the true cluster center is modeled as a Gamma distribution with a characteristic length scale  $R_{\text{mis}} = \tau R_\lambda$ , where  $R_\lambda$  is the cluster radius assigned by redMaPPer, and  $\tau = 0.17 \pm 0.04$ . While the X-ray matched clusters are strongly biased to high richness, the authors do not find a significant richness dependence of their results.

**Photometric redshift estimation:** The DES Y1 redMaPPer photometric redshifts are unbiased at the  $|\Delta z| \leq 0.003$  level, and have a median photometric redshift scatter  $\sigma_z/(1+z) \approx 0.006$  [see Fig. 3 in [15]]. The photometric redshift uncertainties are estimated directly from the photometric data, and are rescaled to match the observed dispersion in spectroscopic cluster redshifts. The photometric redshift errors are both redshift and

richness dependent. The redshift dependence is modeled using a polynomial of order ten, with the coefficients for the polynomial fit independently for each richness bin.

Here, we assume the photometric cluster redshifts are unbiased, and we assume a perfect characterization of the photometric redshift scatter. That is, we do not marginalize over our uncertainty in the scatter in the photometric cluster redshifts. In light of other sources of systematic uncertainty in our analysis—in particular source photometric redshift uncertainties—we are confident that this approximation is sufficient.

Assigning a mass proxy (Richness estimation): If richness is a good mass proxy, then richer clusters should be more massive. As evidenced by [15], this is indeed the case, with the mean mass of galaxy clusters scaling as  $\approx \lambda^{1.3}$ . [24] demonstrated that the redMaPPer richness was the lowest scatter optical mass tracer among those available at the time of that study. Nevertheless, the scatter in mass at fixed richness for redMaPPer clusters is large. Moreover, because of the coarse line-of-sight resolution achievable with broadband photometric survey data, photometric cluster catalogs such as redMaPPer will be susceptible to projection effects [e.g., [31]]. Indeed, there is now ample observational evidence confirming this expectation [32–34]. As emphasized by [35], a detailed quantitative characterization of the impact of projection effects is necessary to derive unbiased cosmological constraints from photometric cluster samples. In this work, we forward-model the impact of projection effects on the DES Y1 cluster sample as described in [36]. This modeling accounts not only for projection effects, but also for the masking of clusters by larger systems during the percolation step of the cluster finding.<sup>11</sup>

### C. Shear Catalog and Associated Systematics

The weak-lensing analysis of [15] relies on the galaxy shape catalogs presented in [37]. In DES Y1, shape measurements have been performed with two independent pipelines, METACALIBRATION [38,39] based on NGMIX [40], and IM3SHAPE [41]. Both codes passed a series of tests that show them to be suitable for cosmological studies. However, for the stacked weak lensing analysis of [15], only the METACALIBRATION shape catalog has been used due its larger effective source density ( $6.28 \text{ arcmin}^{-2}$ ). METACALIBRATION measures shapes by simultaneously fitting the galaxy images in the  $r$ ,  $i$ ,  $z$  bands with a 2D Gaussian model convolved with the point-spread functions (PSF) appropriate to each exposure.

Galaxy shape estimators are subject to various sources of systematic errors. For a stacked shear analysis, the dominant source of uncertainty is a multiplicative bias, i.e., an over- or underestimation of gravitational shear as inferred

from the mean tangential ellipticity of lensed galaxies. METACALIBRATION uses a self-calibration technique to debias shear estimates [37]. Specifically, each galaxy image is deconvolved from the estimated PSF, and a small positive and negative shear is applied to the two ellipticity components of the deconvolved image. The resulting images are then convolved once again with a symmetrized version of the PSF, and an ellipticity is estimated for these new images. This procedure allows one to estimate the response of the shape measurement to gravitational shear from the images themselves. An analogous technique is employed to calibrate shear biases due to selection effects. This involves measuring the mean response of the ellipticities to the selection, and then repeating the selections on quantities measured on artificially sheared images. The effectiveness of the METACALIBRATION self-calibration has been addressed in [37] by means of simulated images generated with the GALSIM package [42] using high-resolution images of the COSMOS field processed to mimic the actual noise and PSFs of the DES Y1 data. From this analysis they obtained a Gaussian prior on the multiplicative bias of  $0.012 \pm 0.013$ , and found no evidence of a significant additive bias term. Among all the sources of multiplicative bias investigated—including errors due to the use of multiepoch data, leakage of stellar objects into the galaxy sample, and errors in the modeling of the PSF—blending is the only component with a net bias. The other sources are consistent with zero bias, although they contribute to the bias uncertainty.

### D. Photometric redshift catalog and associated systematics

Photometric redshifts of source galaxies were estimated using the template-based BPZ algorithm [43,44]. Systematic uncertainties in the recovered redshifts were calibrated in a variety of different ways, including cross-matching to COSMOS galaxies, cross-correlation redshifts [45,46], and through the redshift dependence of the shear signal of foreground galaxies of known redshift [47]. The former two were combined in [48] to arrive at the final systematic error budget for the source photometric redshifts. We emphasize that all three methods resulted in mutually consistent calibrations.

The results of [48] do not directly translate into a calibration of the systematic error associated with photometric redshift estimates in the cluster mass calibration analysis because of differences in how the data are used. Specifically, rather than relying on a tomographic analysis of source galaxies, the cluster mass calibration effort in [15] rescaled the shear signal of each galaxy into the corresponding density contrast variable  $\Delta\Sigma$ . This allowed us to trivially combine the lensing signal of all sources to construct an estimate of the excess surface density profile ( $\Delta\Sigma$ ) of the clusters. [15] used the same COSMOS-matching algorithm of [48] to calibrate the systematic

<sup>11</sup>Percolation refers to removing from the candidate cluster member list galaxies that were blended into richer systems along the line of sight.

uncertainty in the amplitude of the recovered weak-lensing profile due to photometric redshift uncertainties. The principal sources of error in this calibration are the cosmic variance associated with the small area of the COSMOS field and uncertainties in connecting the COSMOS measurements to the source galaxy sample, which result in a 2% systematic uncertainty in the amplitude of  $\Delta\Sigma$ . Here, we make the conservative assumption that this uncertainty is perfectly correlated across all cluster redshifts. The resulting systematic uncertainty in the amplitude of the mass-richness relation of redMaPPer clusters from this effect is 2.6%.

### III. DATA VECTOR AND ERROR BUDGET

The DES Y1 data vector for the cluster abundance analysis comprises:

- (1) the number of galaxy clusters in bins of richness and redshift, and
  - (2) the average mass of the galaxy clusters in said bins.
- We detail below how the data vectors and the associated covariance matrices are constructed, and characterize the associated sources of systematic uncertainty.

#### A. Cluster abundances and uncertainties

We bin the galaxy clusters in three redshift bins spanning the range  $z \in [0.2, 0.65]$  and four richness bins spanning the range  $\lambda \in [20, \infty]$ . The richness selection threshold aims to avoid large fractional uncertainties in cluster richness due to Poisson sampling while the redshift range sampled is driven by the available photometric data: our bluest filter is  $g$ , which restricts our analysis to redshifts  $z \geq 0.2$ , while the depth of the data is such that there are few clusters past  $z = 0.65$ . Table I collates the number of galaxy clusters in each of our richness and redshift bins, as labeled. The binning scheme employed in this work is driven by the weak-lensing analysis of [15], which necessitates somewhat broad bins to achieve high signal-to-noise measurements of the weak-lensing profile of the galaxy clusters. A by-product of this choice is that the number of galaxy clusters in each bin is large; our least populated bin contains 91 galaxy clusters.

The uncertainty in the cluster abundance is modeled as the sum of a Poisson component, a sample variance contribution associated with the unknown density contrast of the DES Y1 survey region as a whole [49,50], and a miscentering component. We note that while the Poisson term of the likelihood is strictly non-Gaussian, the high occupancy number of all of our bins ensures that the Gaussian approximation to the Poisson likelihood is a good approximation.

Sample variance is calculated using the technique of [49]. Briefly, the number density fluctuations in the cluster sample takes the form  $\delta_N = b\delta_V$ , where  $b$  is the bias of the clusters in a given richness/redshift bin, and  $\delta_V$  is the mean matter fluctuation within the appropriate DES Y1 survey volume (there is one such random variable for each redshift bin). The cluster bias as a function of mass is calculated using the fitting formula of [51]. The survey mask is approximated as spherically symmetric about the azimuthal axis. In conjunction with this mask, the redshift intervals for each of the bins defines a survey volume, and  $\delta_V$  is the volume-averaged density contrast  $\delta$ . The associated covariance can be readily calculated in terms of the linear matter power spectrum. We also account for the covariance between neighboring redshift bins. For additional details, we refer the reader to Appendix A in [19]. Our covariance matrix is explicitly model dependent: we compute both Poisson and sample variance contributions at each point in the chain, and we account for the determinant term of the covariance matrix in the likelihood. We have verified that holding the covariance matrix fixed results in nearly identical posteriors. At high richness, the Poisson contribution dominates, with sample variance becoming increasingly important at low richness [49].

Cluster miscentering tends to bias low our richness estimates and induces covariance among neighboring richness bins [e.g., [30]]. Rather than forward modeling this effect we directly correct our observed data vector for it. The correction and the covariance matrix associated with miscentering are estimated as follows: starting from a halo catalog, we assign richness to each halo according to the model of [36]. We then randomly miscenter every halo in the catalog following the miscentering model of [30], and recompute the cluster abundance data vector.

TABLE I. Number of galaxy clusters in the DES Y1 redMaPPer catalog for each richness and redshift bin. Each entry takes the form  $N(N) \pm \Delta N_{\text{stat}} \pm \Delta N_{\text{sys}}$ . The numbers between parenthesis correspond to the number counts corrected for the miscentering bias factors (see Sec. III A). The first error bar corresponds to the statistical uncertainty in the number of galaxy clusters in that bin, and is the sum of a Poisson and a sample variance term. The systematic error is due to miscentering errors in the redMaPPer catalog (see text for details).

$\lambda$	$z \in [0.2, 0.35)$	$z \in [0.35, 0.5)$	$z \in [0.5, 0.65)$
[20, 30)	762 (785.1) $\pm 54.9 \pm 8.2$	1549 (1596.0) $\pm 68.2 \pm 16.6$	1612 (1660.9) $\pm 67.4 \pm 17.3$
[30, 45)	376 (388.3) $\pm 32.1 \pm 4.5$	672 (694.0) $\pm 38.2 \pm 8.0$	687 (709.5) $\pm 36.9 \pm 8.1$
[45, 60)	123 (127.2) $\pm 15.2 \pm 1.6$	187 (193.4) $\pm 17.8 \pm 2.4$	205 (212.0) $\pm 17.1 \pm 2.7$
[60, $\infty$ )	91 (93.9) $\pm 14.0 \pm 1.3$	148 (151.7) $\pm 15.7 \pm 2.2$	92 (94.9) $\pm 14.2 \pm 1.4$



The procedure is iterated  $10^3$  times, and we use these realizations to derive the correction factors—obtained as the mean of the ratios between the number counts in richness/redshift bins including or not the miscentering effect—and the corresponding covariance matrix. The uncertainty associated with cluster miscentering in the abundance function ( $\approx 1.0$ – $1.5\%$ ) is subdominant to the Poisson and sample variance contributions in all richness and redshift bins (see Table I). Note that miscentering only mixes neighboring richness bins at the same redshift; there is no covariance between different redshift bins due to miscentering.

### B. Cluster masses and uncertainties

The mean mass of the galaxy clusters in each richness and redshift bin is estimated through a stacked weak-lensing analysis [15]. Briefly, we use the DES Y1 METACALIBRATION shear catalog [38,39] to estimate the shear for each cluster–source pair. This shear is turned into an estimate of the projected mass-density contrast  $\Delta\Sigma$  using the inverse critical surface density  $\Sigma_{\text{crit}}^{-1}$ . The latter depends on both the lens and source redshifts. For the source redshift, we use the redshift probability distribution for the source as estimated using the BPZ code [43]. The uncertainty in the overall lensing amplitude ( $\langle \Sigma_{\text{crit}}^{-1} \rangle$ ) is calibrated by matching the sources in color-magnitude space to COSMOS galaxies with 30-band photo- $z$ s [52]. In addition, we evaluate the correction to the weak-lensing profiles due to the contamination of the source catalog by cluster members (boost factor) by measuring how interlopers distort the photometric redshift distribution of the source catalog toward the cluster cores. For details, we refer the reader to [53] (see also [54]). The statistical uncertainties of the recovered weak-lensing profiles are characterized using a semianalytic covariance matrix that is validated through comparisons to jackknife estimates of the variance. The covariance matrices account for shape noise, cosmic variance, scatter in the richness-mass relation, scatter in the concentration-mass relation, and scatter in halo ellipticities [15,55]. The covariance matrix on the boost factor profiles are jackknife estimates, but these uncertainties have a negligible impact on the mass posteriors.

We simultaneously fit the recovered weak lensing  $\Delta\Sigma(R|M)$  profile along with the corresponding boost factor

data to arrive at the final posteriors for the mean mass. The theory prediction for  $\Delta\Sigma(R|M)$  is obtained by projecting an analytic model of the halo-mass correlation function. In our fit we only consider data in the radial range  $R \in [0.2, 30]$  Mpc. For each redshift and richness bin considered we vary both the halo concentration and halo mass. Model biases due to our choice of analytic model and the selection effect correction adopted in the *unblinded* analysis are calibrated using numerical simulations. For further details, we refer the reader to section 5.4 of [15] and Appendix D. Table II collects the mean mass estimates and associated errors adopted in the *unblinded* analysis.

We note that the lensing profile  $\Delta\Sigma(R)$  from the data requires an assumed cosmological model to transform angular separations into radial distances and to transform redshifts into angular diameter distances. In addition, the two-halo term of the weak-lensing profile requires that we specify the clustering amplitude of the dark matter. Within the context of a flat  $\Lambda$ CDM cosmological model, this implies that the recovered weak-lensing masses are sensitive to the matter density parameters  $\Omega_m$ , the Hubble parameter  $h$ , and the clustering-amplitude parameter  $\sigma_8$ . The Hubble-parameter dependence can be readily absorbed into the masses by quoting masses in units of  $h^{-1} M_\odot$ . We approximate the dependence of the recovered masses as linear in  $\Omega_m$  and  $\ln(10^{10} A_s)$ . The coefficients of this dependence are evaluated numerically by computing the best-fit masses along a grid of values in  $\ln(10^{10} A_s)$  and  $\Omega_m$ , and fitting the resulting data in each bin with a line. The mean slopes obtained with this procedure are:  $d \log(M)/d\Omega_m = -0.40$  and  $d \log(M)/d \ln(10^{10} A_s) = -0.015$ . We have verified that this approximation is accurate at better than the 2% level in each bin, easily sufficient for our purposes (see Table II). When iterating over the cosmological parameters in our analysis we explicitly account for the above cosmological dependence using this linear approximation.

### C. Systematic error budget

Cluster cosmology has long been limited by systematic uncertainties in cluster mass calibration. This remains true today, and will likely remain so for the foreseeable future. We summarize the observational systematics that we have

TABLE II. Mean mass estimates for DES Y1 redMaPPer galaxy clusters in each redshift bin. The reported quantities are  $\log_{10}(M)$  where masses are defined using a 200-mean overdensity criterion ( $M_{200\text{ m}}$ ). The masses are measured in  $h^{-1} M_\odot$  and include the selection effect correction discussed in Appendix D. The first error bar refers to the statistical error in the recovered mass, while the second error bar corresponds to the systematic uncertainty.

$\lambda$	$z \in [0.2, 0.35)$	$z \in [0.35, 0.5)$	$z \in [0.5, 0.65)$
[20, 30)	$14.036 \pm 0.032 \pm 0.045$	$14.007 \pm 0.033 \pm 0.056$	$13.929 \pm 0.048 \pm 0.072$
[30, 45)	$14.323 \pm 0.031 \pm 0.051$	$14.291 \pm 0.031 \pm 0.061$	$14.301 \pm 0.041 \pm 0.086$
[45, 60)	$14.454 \pm 0.044 \pm 0.050$	$14.488 \pm 0.044 \pm 0.065$	$14.493 \pm 0.056 \pm 0.068$
[60, $\infty$ )	$14.758 \pm 0.038 \pm 0.052$	$14.744 \pm 0.038 \pm 0.052$	$14.724 \pm 0.061 \pm 0.069$



accounted for in our analysis. Where quoted, the numbers refer to the uncertainty in the amplitude of the mass–richness relation, and are taken directly from Table 6 in [15], except as noted below. Multiplicative shear and photometric redshift biases are assumed to be perfectly correlated across all richness and redshift bins. Centering is not assumed to be perfectly correlated across all bins. The systematic errors we have accounted for are:

- (1) Multiplicative shear bias: 1.7% Gaussian ([37], see Sec. II C).
- (2) Photometric redshift bias of the source galaxy population: 2.6% Gaussian [see Sec. 4.3 of [15]].
- (3) Cluster centering:  $\leq 1\%$  [see section 5.2 of [15]]. We forward model the impact of cluster miscentering on the weak-lensing profile, marginalizing over the priors derived by [30] and von der Linden *et al.* (in preparation).
- (4) Modeling systematics: 2% Gaussian [see section 5.4 of [15]]. Inaccuracies in our model of the halo-mass correlation function result in biased mass inferences from the weak lensing data. These biases and their uncertainty are calibrated using numerical simulations.
- (5) Selection effect bias: Systematics which introduce correlation between cluster richness and lensing signal could bias our mass estimates. In [15] we accounted for such bias using an analytical estimate of the impact of halo triaxiality and projection effects on weak lensing mass measurements (see their Sec. 5.4.2). These estimates proved to be significantly smaller than our own, more recent determination using numerical simulations (see Appendix D for details). This simulation analysis lowers the recovered weak-lensing masses in a richness and redshift dependent way, with typical shifts being  $\approx 20\%$ – $30\%$ . The analysis presented in the main text of the paper (*unblinded* analysis) adopts the selection effect corrections derived in Appendix D. We conservatively assume the correction to be uncertain at half its amplitude, leading to an  $\approx 13\%$  systematic uncertainty on mass. This uncertainty accounts for  $\approx 60\%$  of our final error budget on the mass estimates.
- (6) Baryonic physics effects: Following [15] we do not include a systematic error related to baryonic effects, which is estimated to be  $\lesssim 3\%$  [56]. This approximation is justified by the  $R > 200$  kpc radial cut applied to the shear data, and the fact that the concentration parameter is allowed to vary independently, with no informative priors, in each richness and redshift bin.

#### IV. THEORETICAL MODEL

Our theoretical model is the same as that described in detail in [19]. For this reason, here we only provide a summary of our method.

The expectation value of the number counts and mean masses of the redMaPPer galaxy clusters in a given richness and redshift bin are given by

$$\langle N \rangle = \int_0^\infty dz^{\text{true}} \int_{z_{\min}}^{z_{\max}} dz^{\text{ob}} \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda^{\text{ob}} \times \langle n | \lambda^{\text{ob}}, z^{\text{true}} \rangle \frac{dV}{dz^{\text{true}}} P(z^{\text{ob}} | z^{\text{true}}) \quad (1)$$

$$\langle M \rangle = \frac{1}{\langle N \rangle} \int_0^\infty dz^{\text{true}} \int_{z_{\min}}^{z_{\max}} dz^{\text{ob}} \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda^{\text{ob}} \times \langle nM | \lambda^{\text{ob}}, z^{\text{true}} \rangle \frac{dV}{dz^{\text{true}}} P(z^{\text{ob}} | z^{\text{true}}). \quad (2)$$

In the above expressions,  $\lambda_{\min}$  and  $\lambda_{\max}$  are edges of the richness bins, while  $z_{\min}$  and  $z_{\max}$  are the edges of the photometric redshift bins. The quantities  $\langle n | \lambda^{\text{ob}}, z^{\text{true}} \rangle$  and  $\langle nM | \lambda^{\text{ob}}, z^{\text{true}} \rangle$  are the comoving space density of clusters and the mass weighted comoving densities, respectively. The term  $dV/dz^{\text{true}}$  is the survey volume per unity redshift. These various quantities are given by

$$\langle n | \lambda^{\text{ob}}, z^{\text{true}} \rangle = \int_0^\infty dM \frac{dn}{dM} P(\lambda^{\text{ob}} | M, z^{\text{true}}) \quad (3)$$

$$\langle nM | \lambda^{\text{ob}}, z^{\text{true}} \rangle = \int_0^\infty dM \frac{dn}{dM} M P(\lambda^{\text{ob}} | M, z^{\text{true}}) \quad (4)$$

$$\frac{dV}{dz} = \Omega_{\text{mask}}(z) c H^{-1}(z) \chi^2(z) \quad (5)$$

where  $\Omega_{\text{mask}}(z)$  is the survey area as a function of redshift,  $H(z)$  is the Hubble parameters as a function of redshift,  $\chi(z)$  is the comoving distance to redshift  $z$ , and  $dn/dM$  is the halo mass function. The above expression assumes a flat cosmology. The survey area is computed as described in [19], and is nearly constant up to redshift  $z = 0.5$ , dropping to  $\approx 50\%$  of the total survey area at  $z \approx 0.63$ . Uncertainties in the survey area as a function of redshift are below 1%, and do not contribute to our error budget.

As noted earlier in Sec. II B, we assume the photometric redshift probability distributions are known. The halo mass function is modeled using the [57] halo mass function, but allowing for power-law deviations that are calibrated using numerical simulations. Specifically, we assume the mass function is specified by

$$\frac{dn}{dM} = \left( \frac{dn}{dM} \right)_{\text{Tinker}} \left[ s \ln \left( \frac{M}{M_*} \right) + q \right]. \quad (6)$$

The parameters  $s$  and  $q$  are fit to the AEMULUS simulations [58], which are also used to characterize the associated uncertainties in the parameters  $s$  and  $q$  (see Table III). Our cosmological posteriors are marginalized over these uncertainties. Moreover, to keep the universality of the Tinker

TABLE III. Model parameters and parameter constraints from the joint analysis of redMaPPer DES Y1 cluster abundance and weak-lensing mass estimates. In the third column we report our model priors: a range indicates a top-hat prior, while  $\mathcal{N}(\mu, \sigma)$  stands for a Gaussian prior with mean  $\mu$  and variance  $\sigma^2$ . The fourth column lists the modes of the 1-d marginalized posterior along with the  $1-\sigma$  errors. Parameters without a quoted value are those for which the marginalized posterior distribution is the same as their prior.

Parameter	Description	Prior	Posterior
$\Omega_m$	Mean matter density	[0.0, 1.0]	$0.179^{+0.031}_{-0.038}$
$\ln(10^{10}A_s)$	Amplitude of the primordial curvature perturbations	[-3.0, 7.0]	$4.21 \pm 0.51$
$\sigma_8$	Amplitude of the matter power spectrum	—	$0.85^{+0.04}_{-0.06}$
$S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$	Cluster normalization condition	—	$0.65^{+0.04}_{-0.04}$
$\log M_{\min} [M_\odot/h]$	Minimum halo mass to form a central galaxy	(10.0, 14.0)	$11.13 \pm 0.18$
$\log M_1 [M_\odot/h]$	Characteristic halo mass to acquire one satellite galaxy	$\log(M_1/M_{\min}) \in [\log(10), \log(30)]$	$12.37 \pm 0.11$
$\alpha$	Power-law index of the richness-mass relation	[0.4, 1.2]	$0.748 \pm 0.045$
$\epsilon$	Power-law index of the redshift evolution of the richness-mass relation	[-5.0, 5.0]	$-0.07 \pm 0.28$
$\sigma_{\text{intr}}$	Intrinsic scatter of the richness-mass relation	[0.1, 0.5]	$< 0.325$
$s$	Slope correction to the halo mass function	$\mathcal{N}(0.047, 0.021)$	—
$q$	Amplitude correction to the halo mass function	$\mathcal{N}(1.027, 0.035)$	—
$h$	Hubble rate	$\mathcal{N}(0.7, 0.1)$	$0.744 \pm 0.075$
$\Omega_b h^2$	Baryon density	$\mathcal{N}(0.02208, 0.00052)$	—
$\Omega_\nu h^2$	Energy density in massive neutrinos	[0.0006, 0.01]	—
$n_s$	Spectral index	[0.87, 1.07]	—

halo mass function in cosmology with massive neutrinos we adopt the prescription of [59] neglecting the density neutrino component in the relation between scale and mass—i.e.,  $M \propto (\rho_{\text{cdm}} + \rho_b)R^3$ —and using only the cold dark matter and baryon power spectrum components to compute the variance of the density field,  $\sigma^2(R)$ .

Several works have estimated the impact of baryonic feedback on total halo masses and, thereby, the halo mass function [e.g., [60–63]]. These works all find that baryonic impact decreases with increasing mass and radial aperture, while the sign and amplitude of the bias depend on the specifics of the feedback model and simulation adopted. For halos with  $M_{200,m} \gtrsim 10^{14} h^{-1} M_\odot$  and the mass definition we adopt, the baryonic redistribution within a halo due to cooling and feedback processes has a negligible impact on the halo mass function ( $\lesssim 5\%$ ). Given the current uncertainty in modeling baryonic effects, and the  $\sim 5\%$  uncertainty on the Tinker mass function already encoded in the model parameters  $s$  and  $q$ , we do not include additional uncertainties due to the impact of baryonic physics on the halo mass function.

The key remaining ingredient is the model for the richness-mass relation  $P(\lambda^{\text{ob}}|M)$ . Our model is described in [36], which was custom built for this analysis. Briefly, the intrinsic richness-mass relation is modeled using a conventional halo model parametrization, with  $\lambda^{\text{true}} = \lambda^{\text{cen}} + \lambda^{\text{sat}}$  where  $\lambda^{\text{cen}}$  and  $\lambda^{\text{sat}}$  are the number of central and satellite galaxies respectively.  $\lambda^{\text{cen}}$  is assumed to be a deterministic function of mass, with  $\lambda^{\text{cen}} = 1$  for  $M \geq M_{\min}$  and  $\lambda^{\text{cen}} = 0$  otherwise.  $\lambda^{\text{sat}}$  is a random variable with an expectation value

$$\langle \lambda^{\text{sat}} | M, z \rangle = \left( \frac{M - M_{\min}}{M_1 - M_{\min}} \right)^\alpha \left( \frac{1 + z}{1 + z_*} \right)^\epsilon \quad (7)$$

where  $M_1$  is the characteristic mass at which a halo of mass  $M$  has on average one satellite galaxy, and the pivot redshift is set equal to the mean redshift of the sample  $z_* = 0.45$ . Note that the above formula ensures that only halos with central galaxies can have satellite galaxies. To allow for super-Poisson halo occupancies at high mass, we model  $P(\lambda^{\text{true}}|M)$  as the convolution of a Poisson and a Gaussian distribution, where the scatter of the latter is simply  $\sigma_{\text{intr}} \langle \lambda^{\text{sat}} | M, z \rangle$ . For numerical reasons, we approximate this convolution using a skew-normal distribution. For details, see [19], particularly Appendix B. We note that because of the Gaussian component of  $P(\lambda^{\text{true}}|M)$ , a large width may result in negative richness values. These are interpreted as a finite probability of having  $P(\lambda^{\text{true}} = 0)$ , where the probability  $P(\lambda^{\text{true}})$  is set to the integral of the Gaussian model below  $\lambda^{\text{true}} = 0$ . In other words, negative  $\lambda^{\text{true}}$  values are considered halos with no satellite galaxies (and therefore no galaxy overdensity). We investigate the sensitivity of our cosmological conclusions to our model for  $P(\lambda^{\text{true}}|M)$  in Sec. V C.

The observed richness  $\lambda^{\text{ob}}$  is a noisy measurement of  $\lambda^{\text{true}}$ . Four distinct sources of noise on  $\lambda^{\text{ob}}$  are: (1) random errors associated with magnitude errors and background subtraction of uncorrelated structures; (2) projection effects; (3) percolation effects and (4) miscentering effects. The modeling of first three effects is the focus of our work in [36]. In that work, we demonstrate that projection effects follow an exponential distribution, while photometric

uncertainties and background subtraction lead to a Gaussian error. Percolation effects modulate the richness of masked halos by a multiplicative factor that is uniformly distributed between 0 and 1, and the fraction of clusters that suffer from percolation effects is a decreasing function of richness. Parameters governing these distributions are determined by a semiempirical method applied to halos in synthetic light-cone maps derived from N-body simulations [64]. DES redMaPPer data is used to calibrate a projection kernel that is used as a weight function applied to the simulated halos. Using sightlines that target halos of specific intrinsic richness and redshift, a weighted sum of the richness of halos along the line of sight is used to estimate the component of  $P(\lambda^{\text{ob}}|\lambda^{\text{true}}, z)$  arising from two-halo and higher spatial correlations. These same simulations are used to calibrate the purely geometric impact of percolation. The photometric and background subtraction noise is measured by injecting artificial clusters in the data. The end result is a calibrated distribution  $P(\lambda^{\text{ob}}|\lambda^{\text{true}}, z)$  describing the impact of observational uncertainties and projection effects on the DES Y1 redMaPPer cluster sample. Further details of this calibration are presented in Appendix A.

At this point we have described all the necessary ingredients for calculating the expectation value of our observable vector. We model the likelihood function as a Gaussian distribution, which requires that we further specify the associated covariance matrix. As described in Sec. III, the covariance matrix for the abundance reflects Poisson, sample variance, and miscentering uncertainties. This covariance matrix is varied in parameter space, and we explicitly account for the term involving the determinant of the covariance matrix in our likelihood function. The covariance matrix for the recovered weak-lensing masses reflects the semianalytic covariance matrix characterizing the weak-lensing data, and explicitly accounts for systematic uncertainties in the recovered weak-lensing masses. All the systematic uncertainties, except the one associated with selection effects, are assumed to be correlated across richness and redshift bins. The lack of covariance in the selection effects correction allows for the selection effects to vary as a function of richness and redshift.

Our analysis assumes no covariance between the number counts and the recovered mean masses in bins. However, it is reasonable to expect that an increase in projections will give rise to both an increase in the number counts, and an increase in the weak-lensing mass, e.g., due to the effects modeled in Appendix D. Improved simulations and synthetic sky catalogs will allow us to simultaneously model coupled systematic effects within the data vector of counts and mean weak lensing masses. However, large, mass-independent positive correlations between the abundance and weak-lensing masses are ruled out as the resulting covariance matrix stops being positive definite.

In particular, assuming the element of the cross-covariance matrix to be given by  $r\sigma_{\text{NC}}\sigma_{\text{MWL}}$ , as  $r$  increases, the determinant of the covariance matrix decreases, eventually becoming negative at  $r \approx 0.15$ . Adopting a “large” mass-independent correlation coefficient (compared to its maximum possible value above) of  $r = 0.125$  has only a minor impact on our cosmological posteriors, and does not impact any of the conclusions in the discussion below.

### A. Model and data summary

We provide a short, bullet-point summary of our data and model below. Our data can be summarized as follows:

- (i) Our data vector is the DES Y1 redMaPPer cluster counts and weak-lensing masses.
- (ii) The covariance matrix of the cluster counts is due to Poisson noise, sample variance, and cluster miscentering.
- (iii) The covariance matrix of the weak-lensing data is dominated by the impact of selection effects on the weak-lensing profile of the galaxy clusters. The next most important contribution is source photometric redshift uncertainties. The remaining uncertainties are cluster miscentering, lensed galaxy source dilution, and multiplicative shear biases.
- (iv) We assume no covariance between cluster counts and weak-lensing masses.

Our model can be summarized as follows:

- (i) Cluster counts are modeled as a convolution of the [57] mass function with a richness-mass relation.
- (ii) We characterize and account for possible deviations from the Tinker mass function using a suite of numerical simulations.
- (iii) The intrinsic richness of a galaxy cluster  $P(\lambda^{\text{true}}|M)$  is a convolution of Poisson noise with a Gaussian scatter of fixed relative width.
- (iv) The impact of projection effects and observational uncertainties is forward modeled in the counts [36]. There are no nuisance parameters associated with this calibration in our likelihood model.
- (v) Based on numerical simulation estimates we do not assign a systematic error budget to the halo mass function due to baryonic feedback.
- (vi) Based on the fact that the concentration parameter is allowed to float independently in each richness/redshift bin used in the stacked weak lensing analysis, we do not assign a systematic error to the recovered weak-lensing masses due to baryonic effects.
- (vii) Systematic biases (and their uncertainties) due to the use of an analytic halo model for the halo-mass correlation function are calibrated using numerical simulations.

Appendix E applies our methodology to a simulated data set in order to validate the cosmological pipeline.



## V. RESULTS

This analysis has been performed blind following the blinding and unblinding protocol outlined in Appendix B. After unblinding, a  $2.3\sigma$  and  $6.7\sigma$  tension in the  $\sigma_8$ – $\Omega_m$  plane was found with DES 3x2pt [20] and *Planck* CMB data [2], as well as a larger than  $3.5\sigma$  tension with BAO measurements [65–67] and supernovae data [68] (see Appendix C for details). In the attempt to trace back the source of the tension, we found two clear but minor bugs, neither of which had a substantial impact on our posteriors. We also discovered the impact of selection effects on weak lensing in simulations was significantly larger than originally expected (see Appendix D), leading us to revise the estimate of the impact of selection effects on the cluster masses. Below, we present the results for the *unblinded* analysis, which include the selection effects bias estimates from Appendix D. These corrections increased the size of the error ellipse from DES Y1 clusters, but, as discussed below, significant tension with *Planck* and DES 3x2pt remains. If not specified otherwise, we assume a flat  $\Lambda$ CDM cosmological model with three degenerate species of massive neutrinos ( $\Lambda$ CDM +  $\sum m_\nu$ ). The parameter

posteriors are estimated using the EMCEE package [69] which implement the affine-invariant Monte Carlo Markov Chain sampler of [70].

### A. Goodness of fit

Figure 2 shows the abundance (left) and weak-lensing masses (right) of the DES Y1 redMaPPer clusters as a function of the cluster richness for three separate redshift bins along with the corresponding best-fit model expectations. The measurements and associated uncertainties are shown as colored boxes, while the dots correspond to the best-fit model from our posteriors. The bottom panel shows the residual between the data and our best-fit model for each of the three redshift bins under consideration, as labeled. For clarity, the points are slightly spread along the richness axis to avoid overcrowding. The  $\chi^2$  of our best-fit model is 22.33.

We assess the goodness of fit by generating 100 realizations of our best-fit model data vectors adopting our best-fit covariance matrix, and fitting each in turn in order to arrive at the distribution of best-fit  $\chi^2$  values of our mock-realizations. The distribution is fit using a  $\chi^2$

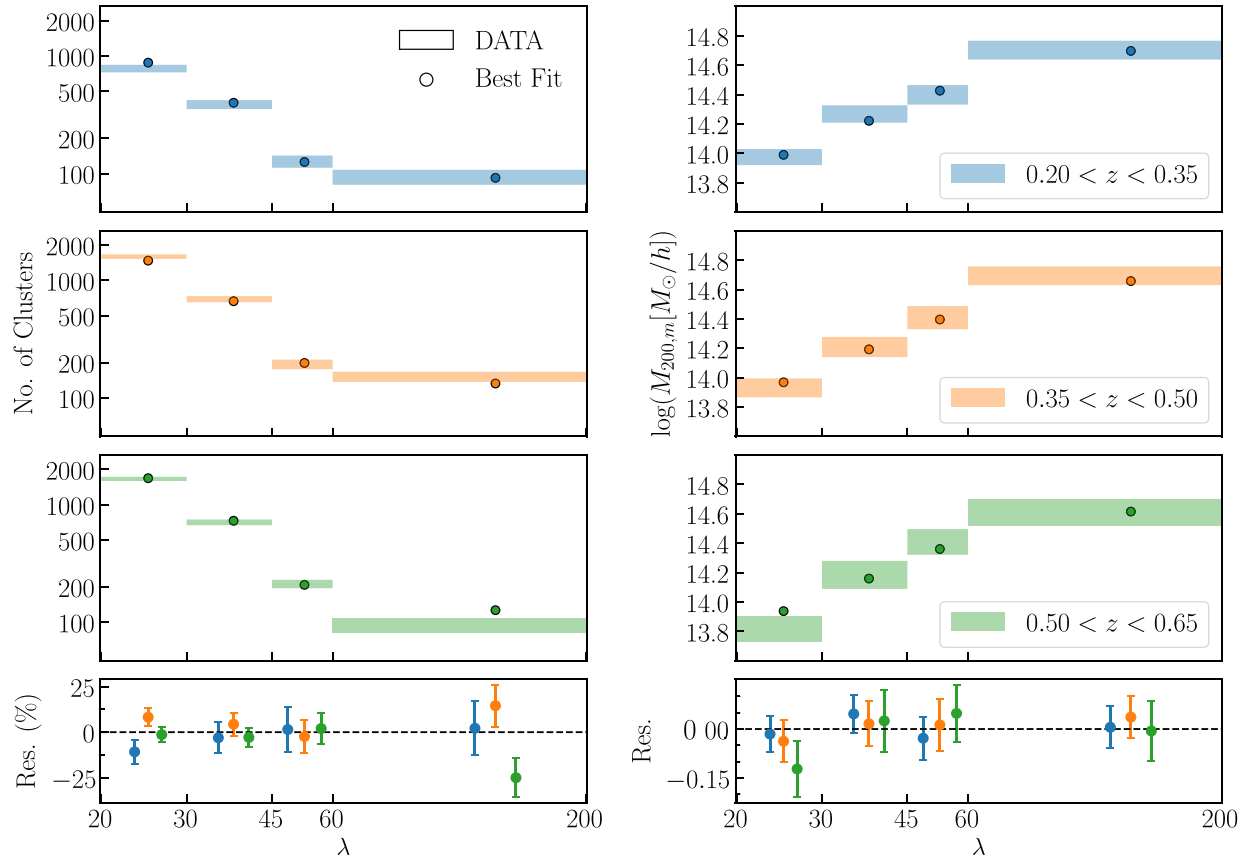


FIG. 2. Observed (shaded areas) and best-fit model (dots) for the cluster number counts (left) and mean cluster masses (right) as a function of richness for each of our three redshift bins. The y extent of the data boxes is given by the square root of the diagonal terms of the covariance matrix. The bottom panel shows the residual between the data and our best-fit model. All points have been slightly displaced along the richness axis to avoid overcrowding.

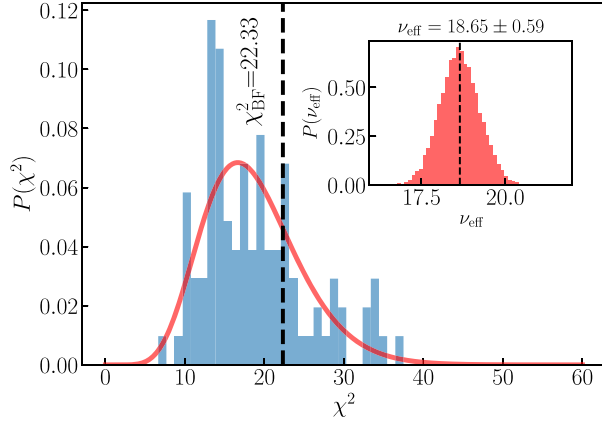


FIG. 3. Goodness-of-fit analysis. The blue histogram shows the distribution of the best-fit  $\chi^2$  values recovered from 100 mock data realizations generated from the best-fit model of the data. The red histogram in the inset plot shows the posterior distribution for the effective number of degrees of freedom obtained by fitting a  $\chi^2$  distribution to the above 100  $\chi^2$  values. The red solid line represents the  $\chi^2$  distribution for the best-fit model ( $\nu_{\text{eff}} = 18.65$ ), while the vertical dashed line corresponds to the  $\chi^2$  value of the data.

distribution, for which we find that the effective number of degrees of freedom is  $\nu_{\text{eff}} = 18.65 \pm 0.60$ . The distribution of  $\chi^2$  values in our simulated data, as well as the  $\chi^2$  value in the real data, is shown in Fig. 3. As evident from the figure, our model is a good fit to the data, with a probability to exceed of 0.25.

### B. Cosmological constraints from DES Y1 cluster data

Figure 4 shows the posteriors of the parameters used to model the DES Y1 cluster cosmology data set. The parameter  $M_{\text{min}}$  is not shown because it is prior dominated. All of our parameters, along with their corresponding priors and posteriors, are summarized in Table III.

The only two cosmological parameters that are not prior dominated in our analysis are  $\sigma_8$  and  $\Omega_m$ . Our posteriors for each of these are  $\sigma_8 = 0.85^{+0.04}_{-0.06}$  and  $\Omega_m = 0.179^{+0.031}_{-0.038}$ . The corresponding cluster normalization condition is  $S_8 = \sigma_8(\Omega_m/0.3)^{0.5} = 0.650 \pm 0.037$ .

In addition, the posterior for the Hubble parameter  $h = 0.744 \pm 0.075$  is slightly improved relative to our prior,  $h = 0.7 \pm 0.1$ . This improvement arises due to the mild sensitivity of number counts and mean cluster masses to  $h$ : a shift of  $h$  tilts the slope of the number counts around the pivot point  $\lambda \simeq 55$  while changing the amplitude of the mean mass–richness relation. Despite the modest degeneracy of  $h$  with  $\Omega_m$  and  $\sigma_8$ , we verified that adopting a flat prior on  $h \in [0.55, 0.90]$  (as in DES Collaboration *et al.* [20]) does not affect the cosmological posteriors of  $\Omega_m$  and  $\sigma_8$ .

We compare our posterior on the parameter  $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$  to that derived from a variety of different weak lensing and cluster abundance experiments in Fig. 5. This figure also compares our posterior in  $S_8$  to that of *Planck* 2016 and *Planck* 2018. Our posterior is clearly lower than all other constraints shown, with the tension in  $S_8$  relative to other low-redshift probes typically ranging from  $1.5\sigma$  to  $2.5\sigma$ . Notably, one of the largest tensions is with respect to the DES Y1 3x2pt analysis, at  $2.9\sigma$ . We note that these tensions in  $S_8$  were only slightly impacted by the post-unblinding corrections we adopted. If we naively combine all nine low-redshift experiments assuming they are mutually independent, the DES Y1 cluster result has a 2% probability of being a statistical fluctuation around their mean. The difference becomes even stronger when considering *Planck* CMB results, for which the significance of the tension with  $S_8$  reaches  $4.0\sigma$ .

Figure 6 compares the 68% and 95% confidence regions in the  $\sigma_8$ – $\Omega_m$  plane derived from DES Y1 cluster data to the DES 3x2pt statistics [20], the *Planck* CMB DR18 [2], a combination of BAO measurements [65–67], Supernovae Pantheon data [68], and cluster counts analyses from [7,9] (respectively WtG and SPT-2500 in the figure). As is evident from the figure, the  $S_8$  tension is due to the low  $\Omega_m$  value preferred by the DES Y1 cluster data set. Specifically, looking at the  $\Omega_m$  subspace, our cluster posterior displays a  $1.7\sigma$  tension with SPT-2500,  $1.8\sigma$  tension with WtG, a  $2.2\sigma$  tension with DES Y1 3x2pt, a  $3.0\sigma$  tension with SN data, a  $3.3\sigma$  tension with BAO, and a  $4.7\sigma$  tension with *Planck* CMB. The corresponding tensions in the  $\sigma_8$ – $\Omega_m$  plane are  $1.1\sigma$  (SPT-2500),  $1.7\sigma$  (WtG),  $2.4\sigma$  (DES 3x2pt) and  $5.6\sigma$  (*Planck*).<sup>12</sup> The fact that all other cosmological probes, including those using the same DES data employed in this work, return significantly higher values for the matter density than ours suggests the presence of unexpected systematics or physics in our analysis. We will comment on the possible origin of this tension in Sec. VI. Due to the inconsistencies between the DES Y1 cluster data and internal and external probes we do not perform any joint analysis of cluster data with other data sets.

One intriguing possibility to consider is whether the tensions seen in Fig. 6 could be reduced within the context of a different cosmological model. We have run chains assuming a  $w\text{CDM} + \sum m_\nu$  model with a flat prior  $w \in [-2, -1/3]$  for the equation of state of the dark energy. We find that these models do not improve the agreement between DES clusters and the remaining data sets.

<sup>12</sup>Here consistency between two data sets  $A$  and  $B$  is established by testing whether the hypothesis  $\mathbf{p}_A - \mathbf{p}_B = 0$  is acceptable [see method ‘3’ in [77]], where  $\mathbf{p}_A$  and  $\mathbf{p}_B$  are the model parameters of interest as constrained by data sets  $A$  and  $B$ , respectively.

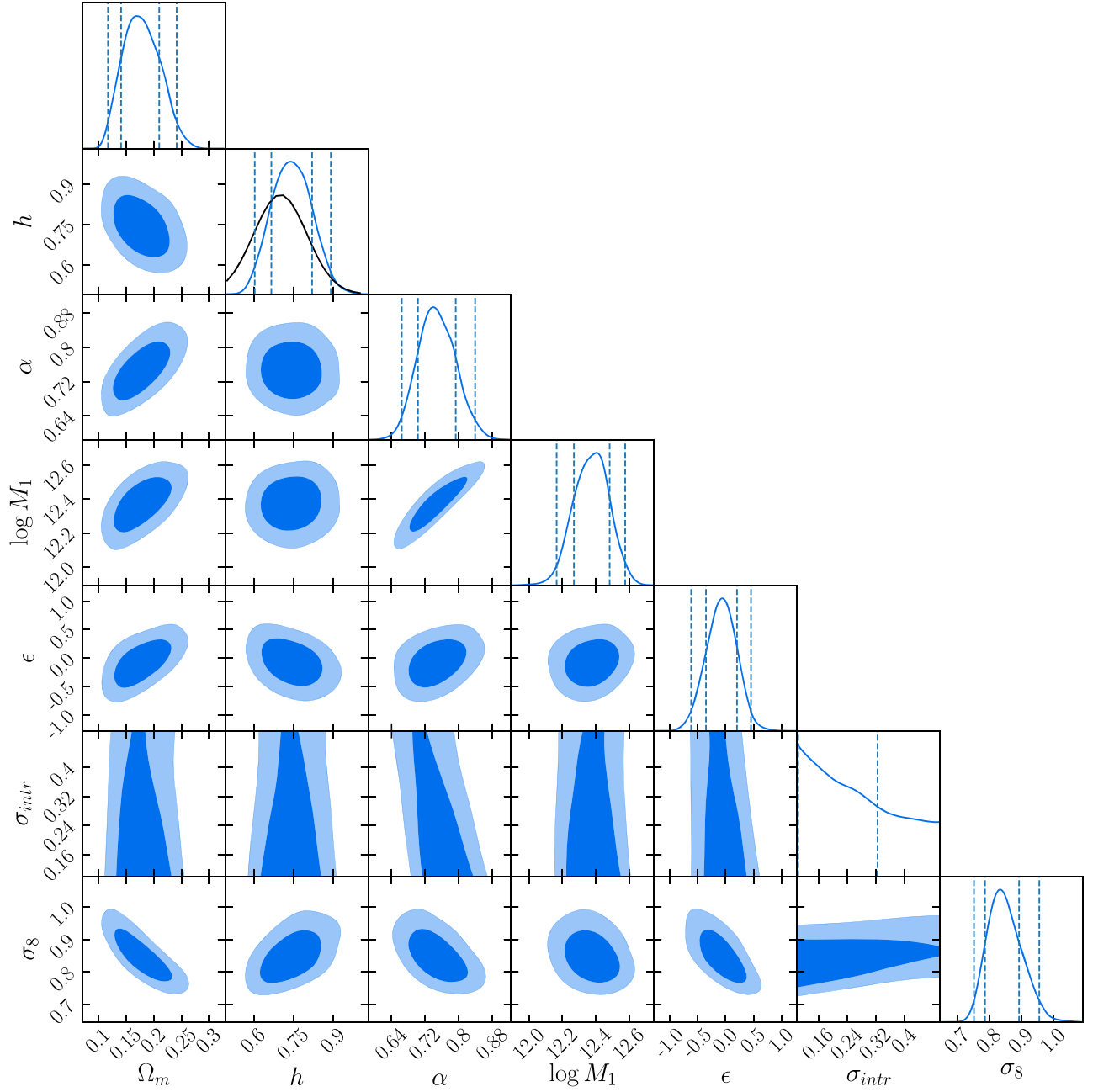


FIG. 4. Marginalized posterior distributions of the fitted parameters. The 2D contours correspond to the 68% and 95% confidence levels of the marginalized posterior distribution. The dashed lines on the diagonal plots correspond respectively to the 2.5th, 16th, 84th and 97.5th percentile of the 1-d posterior distributions. The black line in the 1-d posterior plot of  $h$  corresponds to the Gaussian prior adopted in the analysis. The description of the model parameters along with their posteriors are listed in Table III. Only parameters that are not prior dominated are shown in the plot.

### C. Robustness tests

Of special interest to us is the robustness of our cosmological posteriors to our choice of theoretical model. To test for robustness we consider three different modifications to our fiducial model for the richness-mass relation, which in turn affect the expectation values for the number counts and mean cluster masses. These are

- (1) A random-point injection model, in which projection effects are estimated assuming clusters are randomly located throughout the sky. This provides a firm lower limit on projection effects. We consider this an extreme model (i.e., we know clusters live in highly clustered regions of the Universe).
- (2) A model with boosted projection effects, in which  $P(\lambda^{\text{ob}}|\lambda^{\text{true}})$  is calibrated doubling the magnitude of



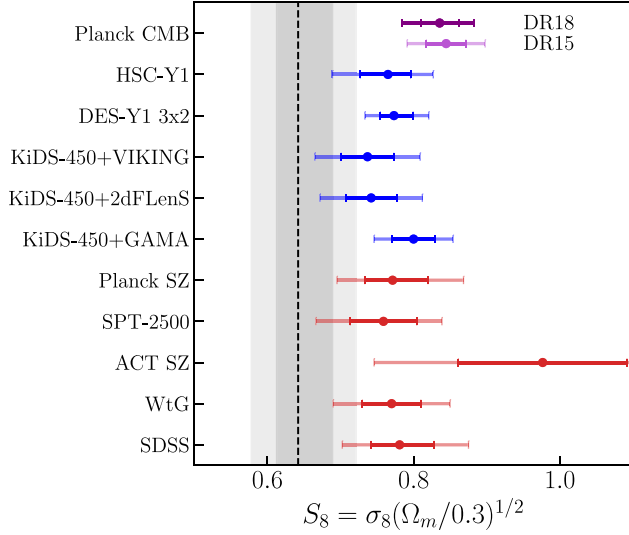


FIG. 5. Comparison of the 68% (dark) and 95% (light) confidence level constraints on  $S_8$  derived from our baseline model (shaded gray area) with other constraints from the literature: red error bars for cluster abundance analyses, blue error bars for weak lensing and galaxy clustering analyses and purple for the CMB constraint. From the bottom to the top: SDSS from [19]; WtG from [7]; ACT SZ from [71] (BBN + H0 + ACTcl(dyn) in the paper); SPT-2500 from [9]; Planck SZ from [72] (CCCP +  $H_0$  + BBN in the paper); KiDS-450 + GAMA from [73]; KiDS-450 + 2dFLenS from [74]; KiDS-450 + VIKING from [75]; DES-Y1  $3 \times 2$  from [20]; HST-Y1 from [11]; Planck CMB from [76] (DR15) and [2] (DR18). Note that all the constraints but those from SDSS, DES-Y1  $3 \times 2$ , HSC-Y1 and Planck CMB have been derived fixing the total neutrino mass either to zero or to 0.06 eV.

projection effects relative to our fiducial model. We expect this model provides an upper limit on the effect that an underestimation of projection effects could have on cosmological posteriors.

(3) A model in which  $P(\lambda^{\text{ob}}|M)$  is a log-normal, the mean richness-mass relation is a power law and the intrinsic scatter is mass dependent; note that in this case we do not include our model for  $P(\lambda^{\text{ob}}|\lambda^{\text{true}})$ , and all the scatter due to observational noise and projection effects is absorbed by the  $\sigma_{\text{intr}}$  parameter. As detailed in Appendix B, these models were selected and tested before unblinding. We thus repeated these tests for the *unblinded* analysis finding consistent effects on the parameter posteriors to those obtained in the *blinded* analysis. Figure 7 shows how our cosmological posteriors of the *unblinded* analysis change for each of these different model assumptions. As noted above, we consider model (i) to be extreme and (ii) to provide a conservative upper limit on the amplitude of projection effects, and use them to define a  $2\sigma$  systematic error in our cosmological parameters associated with the projection-effect calibration. That is, we estimate the systematic uncertainty in our cosmological posteriors as half the difference between the recovered

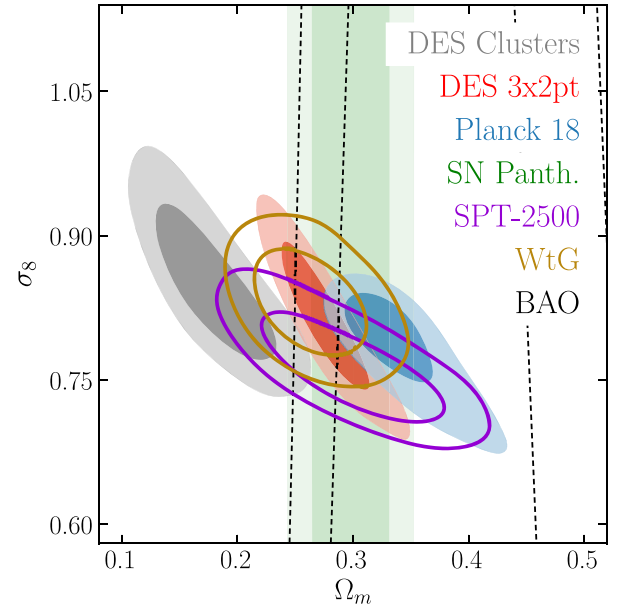


FIG. 6. Comparison of the 68% and 95% confidence contours in the  $\sigma_8$ - $\Omega_m$  plane derived from DES Y1 cluster counts and weak-lensing mass calibration (gray contours) with other constraints from the literature: BAO from the combination of data from Six Degree Field Galaxy Survey [6dF [65]], the SDSS DR 7 Main galaxy sample [66], and the Baryon Oscillation Spectroscopic Survey [BOSS [67]] (black dashed lines); Supernovae Pantheon [68] (green contours); DES-Y1  $3 \times 2$  from [20] (red contours); Planck CMB from [2] (blue contours); SPT-2500 from [9] (violet contours); WtG from [7] (gold contours).

parameters in these models and our fiducial model. These systematic errors are negligible compared to our posteriors, and will therefore be ignored from this point on.

Similarly, the central values of our cosmological posteriors when using model (iii) are within the one-sigma posterior of our reference model. We include this model here for comparison purposes, since previous analyses have relied on power-law log-normal models [e.g., [10,78]].

Appendix D details further tests of the parametrization of the richness-mass relation performed after unblinding. The summary of those results is consistent with our conclusions above: the adopted form of the richness-mass relation does not have a large impact on the cosmological posteriors derived from our analyses.

#### D. Constraints on the richness-mass relation

Figure 8 shows the posterior of the richness-mass relation of the DES Y1 redMaPPer galaxy clusters. The left panel shows the expectation value of the richness-mass relation,  $\langle \lambda^{\text{ob}}|M \rangle$  at the mean sample redshift  $z = 0.45$ . The central panel shows the variance in richness at fixed mass,  $\text{Var}(\lambda^{\text{ob}}|M)$ , again at the mean sample redshift. It is important to emphasize that the shape of the variance as a function of mass is intrinsic to our fiducial model: while

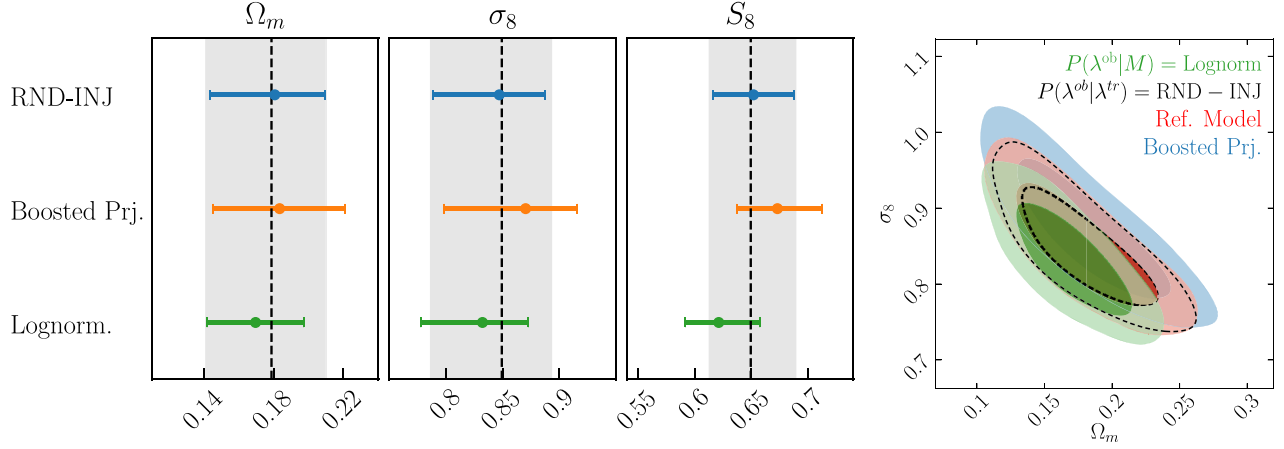


FIG. 7. *Left panel:* Comparison of the 68% confidence regions for  $\Omega_m$ ,  $\sigma_8$  and  $S_8$  derived from different model assumptions. The shaded area corresponds to the constraints derived using our reference model, while the dashed vertical line is the mean of our fiducial posterior. The model assumptions we consider are, from top to bottom, the random injection, boosted projection, and lognormal models described in Sec. V D. *Right:* Same as *left* panel in the  $\Omega_m$ - $\sigma_8$  plane.

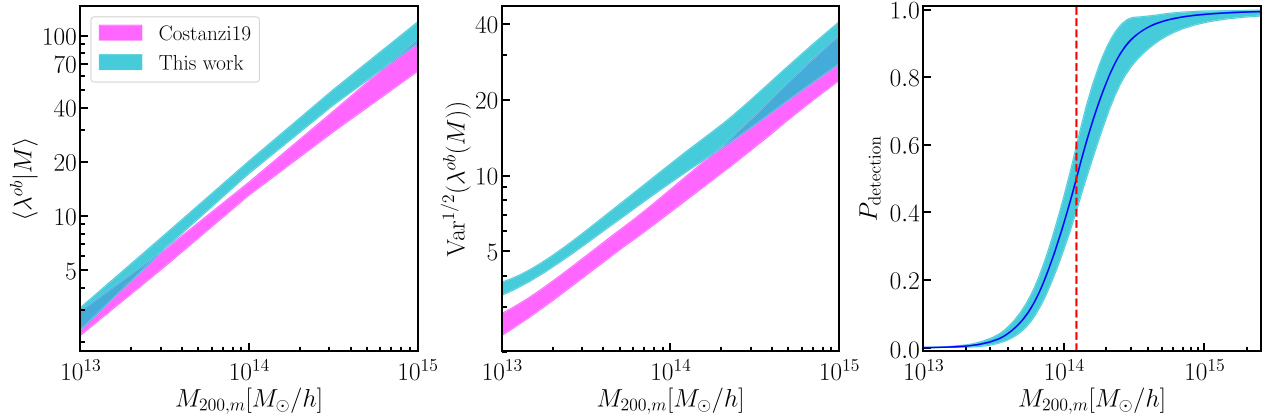


FIG. 8. Observable-mass relation and mass-selection function of the redMaPPer DES Y1 catalog assuming our reference richness-mass relation model [Eq. (7)] at the mean sample redshift  $z = 0.45$ . Left panel: Expectation value for the observed richness as a function of mass. Central panel: Scatter of  $\lambda^{\text{ob}} - \text{Var}^{1/2}(\lambda^{\text{ob}}|M)$ —as a function of mass. Right panel: Detection probability as a function of cluster mass. The dashed vertical line corresponds to the mass at which the detection probability is 50% ( $\log M_{50\%} = 14.09 [M_\odot/h]$ ). The blue area corresponds to the 68% confidence interval derived for the different quantities in this work. For comparison, the results of [19] (magenta) for the redMaPPer SDSS catalog are shown in the two left panels. All results are corrected for the systematic richness offset between the SDSS and DES catalogs, and account for the expected redshift evolution between  $z = 0.22$  (SDSS) and  $z = 0.45$  (DES) as determined by the best-fit model to the DES data.

we have a single scatter parameter  $\sigma_{\text{intr}}$ , which is mass independent, our model for both the intrinsic richness-mass relation and projection effects results in a mass-dependent variance. Finally, the right panel of Fig. 8 shows the probability that redMaPPer will detect a halo of mass  $M$  as a cluster with more than 20 galaxies. The mass at which the detection probability is 50% is  $M_{200\text{ m}} = 1.2 \times 10^{14} h^{-1} M_\odot$ .

Figure 8 also compares our posteriors to those of our analysis of the SDSS redMaPPer cluster sample [19]. For the purposes of this comparison, we cross match low-redshift DES clusters with SDSS clusters, and correct the SDSS richnesses for the systematic richness offset of 0.93

between SDSS and DES [Eq. (67) in [15]]. Further, we correct our SDSS result for the expected redshift evolution from  $z = 0.23$ —the mean redshift of the SDSS redMaPPer clusters—to our chosen pivot point of  $z = 0.45$  using the best-fit value for the evolution parameter  $\epsilon$  from the DES chain. While the slopes of the richness-mass relations are in agreement between the two analyses, the DES data prefers a larger value for the amplitude. This difference is explained by the selection effect bias correction applied to the weak-lensing mass estimates (see Appendix D): while the mass estimates in [15] were consistent with those of SDSS redMaPPer clusters [18], our selection effect correction lowered the DES Y1 masses by  $\sim 20\%$  relative to our

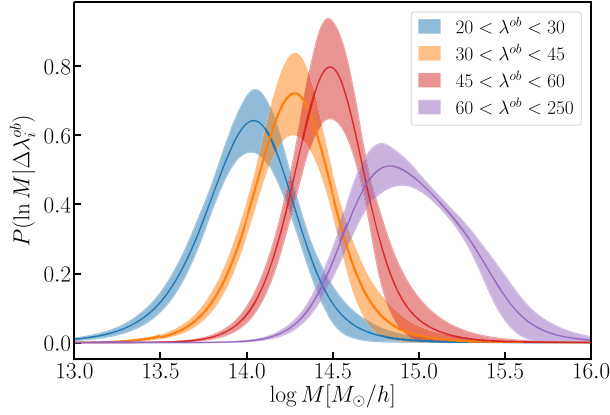


FIG. 9. Distribution of halo mass for clusters in each of the four richness bins employed in this work at median redshift  $z = 0.45$ , as labeled. The width of the bands correspond to the 68% confidence interval of the distribution as sampled from our posterior.

analysis in [15]. By the same token, the variance as a function of mass is similar between the two analyses, but shifted to lower masses in this work because of the selection effects correction. We note, however, that the selection effects characterized in this work should also impact the SDSS constraints. That is, we expect the SDSS richness-mass relation shown above to be biased low by  $\approx 15\%$ .

Figure 9 shows the mass distribution for each of our four richness bins at a redshift  $z = 0.45$ , as constrained through our posteriors. Integrating over these distributions, we can recover the mean mass of the redMaPPer galaxy clusters of a given richness. This mean mass is shown with a *blue* band in Fig. 13. From the combination of DES Y1 cluster counts and weak-lensing mass estimates we constrain the mean mass at the pivot richness  $\lambda^{\text{ob}} = 40$  to  $\log \langle M | \lambda^{\text{ob}} \rangle = 14.252 \pm 0.026$ . As before, the selection effect bias correction applied in this work lowered our masses by  $\sim 20\%$ , leading to a mismatch between our results and that presented in [15]:  $\log(M_0[M_\odot/h]) = 14.334 \pm 0.022$ . Remarkably, the  $\approx 6\%$  precision in the posterior masses is similar to the uncertainty quoted in [15], despite the large systematic uncertainty we have added to the weak lensing masses. This demonstrates that the inclusion of cluster count data offsets the factor of  $\sim 2$  larger uncertainty in mass due to the uncertain calibration of selection effects in our final results. However, the calibration of the scaling relation through number counts data is made at the expense of more relaxed cosmological constraints. For the same reason, this posterior would likely relax in extended cosmological models such as  $w\text{CDM}$ .

## VI. DISCUSSION

### A. What drives the tension between DES clusters and other probes?

The internal consistency of the other DES probes, along with their consistency with external cosmological probes,

rule out the possibility that the tension observed with the DES Y1 clusters data is driven by observational systematics affecting the DES data (e.g., photometry or shear calibration). Thus, the tension between our results and other cosmological probes provides strong evidence that at least one aspect of our theoretical model is incorrect: either the cosmological model assumed is wrong ( $\Lambda\text{CDM} + \sum m_\nu$  and  $w\text{CDM} + \sum m_\nu$ ), our interpretation of the stacked weak lensing signal as mean cluster mass is incorrect, or our understanding of the richness-mass relation and/or selection function is flawed. The interpretation of our results as evidence for the first is unlikely: it would require our analysis to be correct, while all other cosmological experiments would need to have large, as of yet undiscovered systematics. Turning to our understanding of the richness-mass relation, we have verified (Sec. V C) that our cosmological conclusions are robust to the form of the richness-mass relation adopted within the uncertainty suggested by numerical simulations and data. As discussed below, while additional observational tests will be critical to further validate it, currently available multiwavelength data already disfavour the possibility that an unmodeled systematic in  $P(\lambda^{\text{ob}}|M)$  could fully account for the bias in our cosmological posteriors. Given the surprisingly large impact of selection effects in simulations, and that these effects have only been calibrated with one set of simulations, it appears likely that it is our understanding of selection effects on the weak-lensing signal where our model fails.

To study possible unmodeled systematics in our data, we separately reanalyze either the number counts or the weak-lensing mass data, adopting as priors the cosmological posteriors derived from the DES 3x2pt analysis [20]. By doing so, we can compare the posteriors of the richness-mass relation derived using each of our two types of cluster observables independently. The result of this exercise is shown in Fig. 10. Green contours are derived from the combination of number counts data and DES Y1 3x2pt priors, while the black dashed contours combine the Y1 3x2pt priors with the cluster mass data only. Also shown in red for comparison are our reference model posteriors obtained from the combined analysis of number counts and weak lensing data.

As expected, in both cases the cosmological posteriors are dominated by the DES Y1 3x2pt priors, while the richness-mass relation parameters are constrained by either the cluster counts or the weak-lensing mass data alone. It is clear from Fig. 10 that the posteriors for the richness-mass relation derived from either of the cluster observables assuming a DES 3x2pt cosmology are only marginally consistent with one another. In particular, the abundance data prefer a steeper slope and a larger normalization for the richness-mass relation compared to the weak lensing data. This is not unexpected: had they been consistent, we would have expected the DES 3x2pt cosmology to be contained



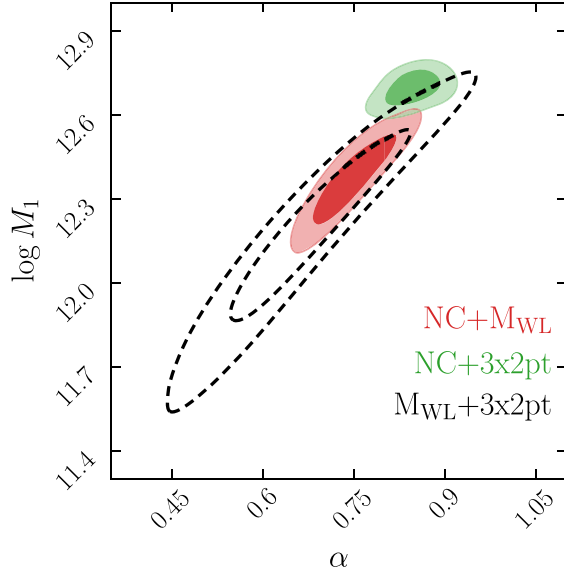


FIG. 10. 68% and 95% confidence contours from the combination of cluster counts data and DES Y1 3x2pt cosmology (green) and the combination of weak-lensing mass estimates and DES Y1 3x2pt cosmology (black). For comparison also shown in *red* our reference model results from the combination of cluster counts and weak lensing data. Not shown in the plot are the  $\sigma_{\text{intr}}$  posteriors since the parameter is not constrained without the inclusion of number counts data ( $M_{\text{WL}} + 3x2pt$ ), whereas we recover the reference model posterior in the NC + 3x2pt case.

within our joint cosmological posterior. The marginal consistency reflects the fact that our posteriors are only marginally consistent ( $2.4\sigma$ ) with the DES 3x2pt cosmology constraints. Interestingly, [79] found a similar trend between the slope preferred by either weak lensing data or cluster abundance when analyzed separately for the first-year HSC data set in a Planck cosmology. However, a direct comparison with our results is not feasible due to the different richness definition and richness-mass relation adopted in their work.

We may now use the posteriors of the richness-mass relation derived using one observable (cluster counts or cluster masses) to predict the complementary observable. This allows us to determine which aspects of the data are driving the tension in Fig. 6. Figure 11 shows the comparison of our data vectors (shaded areas) with our two predictions based on the complementary data set combined with DES 3x2pt priors (filled circles with error bars).

We see that the assumption that our recovered cluster masses and 3x2pt cosmology are correct implies that the redMaPPer catalog is highly incomplete. Specifically, redMaPPer should be  $\sim 50\%$  incomplete at low richness, and between 10%–40% incomplete in the highest richness bin. The redMaPPer catalogs have been extensively vetted over the years, and such a large incompleteness, especially at high richness, is unlikely. For instance, 100% of the SPT and Planck SZ clusters within the DES Y1 footprint and

below redshift 0.65 are detected by redMaPPer. Extensive cross checks with both SPT cluster samples at  $z > 0.25$  [80,81] and x-ray cluster samples at  $0.1 < z < 0.35$  [82] have so far failed to identify a single instance of a clear nondetection of a galaxy cluster due to redMaPPer algorithmic failures. In short, while there is still some room for a small fraction of undetected clusters at low richness, the level of incompleteness in the number counts required at  $\lambda \gtrsim 40$  by our weak lensing cluster masses in a 3x2pt cosmology is unfeasible.

The right panels of Fig. 11 compare the cluster masses predicted by the cluster counts assuming a 3x2pt cosmology to the masses estimated using weak lensing. We find that the weak-lensing masses are low relative to the predicted masses based on the cluster number counts using the 3x2pt cosmology, with the difference ranging from  $\sim 10\%$  percent in the highest richness bins to  $\sim 30\text{--}40\%$  in the lowest richness bins. In other words, the slope of the recovered mass-richness relation from our weak lensing analysis appears to be biased high, a point to which we will return below.

With the exception of our lowest richness bins, the difference between our predicted and observed weak-lensing masses can be reconciled within the systematic uncertainty associated with the selection effects corrections. It is interesting that interpreting the tension in terms of selection effect bias requires lowering the amplitude of the selection effect correction derived in Appendix D to a level comparable to our pre-unblinding analytical estimates. This is shown most clearly in Fig. 12, in which we compare the correction to the “raw” weak-lensing masses necessary to reconcile the weak-lensing data with the number counts within the context of a DES Y1 3x2pt cosmology (cyan bars) with the selection effect correction applied to the data (orange bars). There are two key takeaways from this figure: (1) the simulation-based estimates of the impact of selection effects appear to overcorrect the weak-lensing masses, with the original analytical estimates being closer to what we would expect given a DES 3x2pt cosmology and the observed cluster counts, and (2) remarkably, a DES 3x2pt cosmology requires that we increase the recovered weak-lensing masses in our lowest richness bins by  $\approx 30\%$  to be consistent with our number counts. The fact that the weak-lensing masses of the low richness clusters are biased low is counter to our *a priori* expectations.

## B. What are possible solutions?

If we interpret our results as due to an offset between the recovered weak lensing masses and true mean cluster masses, Fig. 12 poses a remarkably difficult challenge. First, in order to match the DES 3x2pt expectation, the resulting bias must be richness dependent. This immediately rules out traditional weak-lensing systematics—e.g., source photometric redshifts and/or multiplicative shear

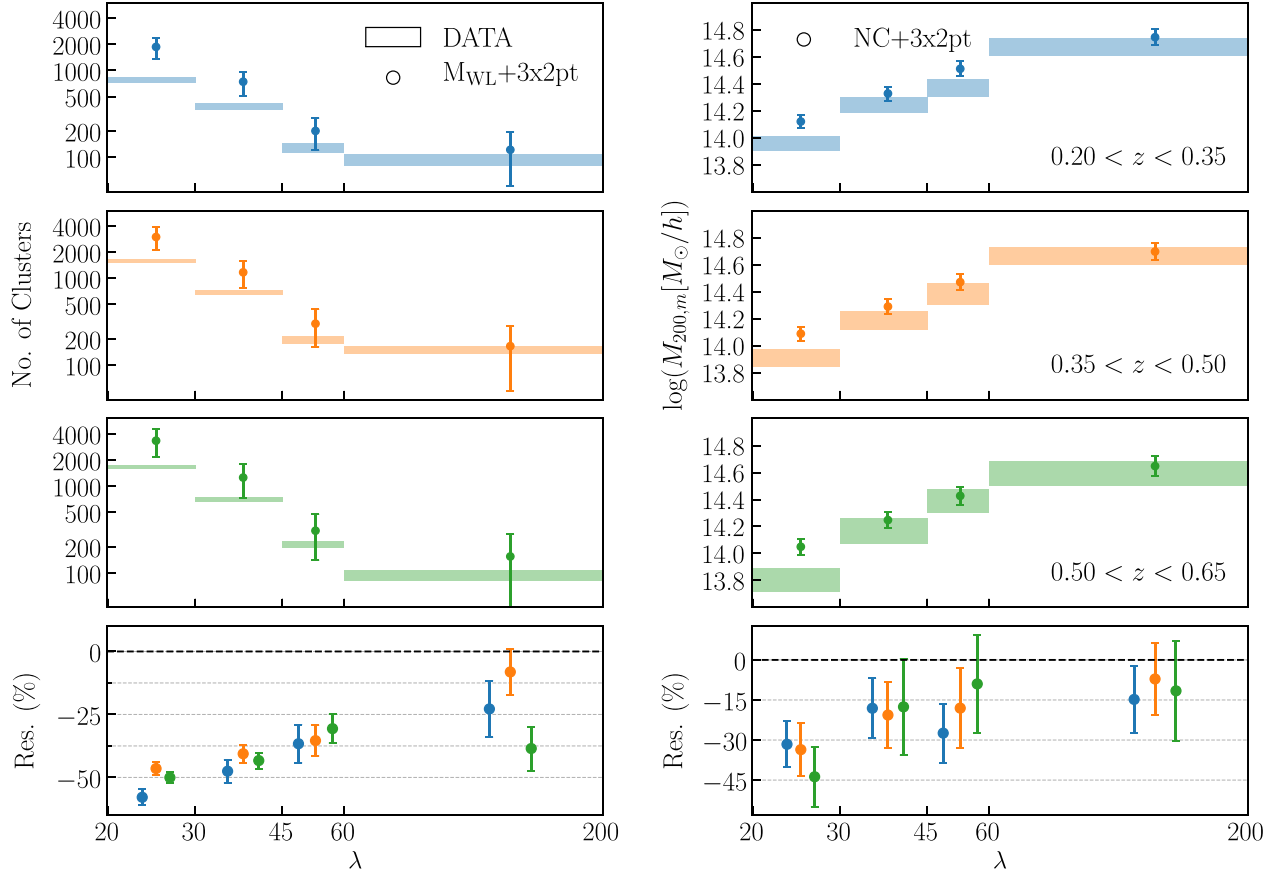


FIG. 11. Comparison of the observed data vectors (shaded areas) with the number counts predicted from the combination of weak-lensing mass estimates and DES Y1 3x2pt cosmology (left panel), and mean masses predicted from the combination of Y1 number counts data and DES Y1 3x2pt cosmological priors (right). The y extent of the shaded areas correspond to the error associated with the data. The error bars on the predicted number counts and mean masses represent one standard deviation of the distribution derived sampling the corresponding MCMC chain. The lower panel shows the percent residual of the predictions to the data vectors, where the error bars refer to data vector uncertainties.

biases—since these systematics give rise to coherent shifts in the recovered masses across all richnesses. It is also worth noting that in addition to our own weak-lensing analysis, [83,84] used CMB lensing signal around DES clusters to determine the amplitude of the mass-richness relation, finding results consistent with our own. This further strengthens the case that the weak-lensing signal is being measured correctly, but that its interpretation in terms of mean true mass is potentially problematic.

Perhaps the biggest challenge that Fig. 12 poses is the fact that while the “raw” weak-lensing masses are biased high at high richness (as expected), at low richness the weak-lensing masses are biased low by a very large amount. Since projection effects and cluster triaxiality tend to boost richness and weak-lensing masses in concert—leading to raw weak-lensing masses that are biased high—Figure 12 suggests that these systematics are incapable of reconciling the weak lensing and abundance data within the context of a DES 3x2pt cosmology.

The above argument assumes that projection effects act primarily as a form of noise that boosts the richness and

weak-lensing masses of existing clusters, but one might wonder whether projection effects are better thought of as creating “false detections” in which “clusters” are really a string-of-pearls type arrangement, with no especially massive halo along the line of sight. One way to think of such projections is as very large non-Gaussian tails in the richness-mass relation toward high richness. From Fig. 7, we see that doubling the amount of projection effects in our galaxy clusters moves our cosmological posteriors toward the DES 3x2pt model. However, a further increase of the amplitude of projection effects will not correspond to an additional relaxation of the tension with DES 3x2pt: the benefit of lowering the predicted mean cluster masses will be counterbalanced by the worse fit to the abundance data due to the predicted larger number of clusters.

More quantitatively, we assess the capability of a large contamination fraction to relieve the tension with 3x2pt as follows: we consider a model in which a fraction  $f_{\text{cont}}$  of the detected clusters is contributed by line-of-sight projections with effectively zero weak-lensing mass. To account for this systematic, we rescale the predicted number counts and

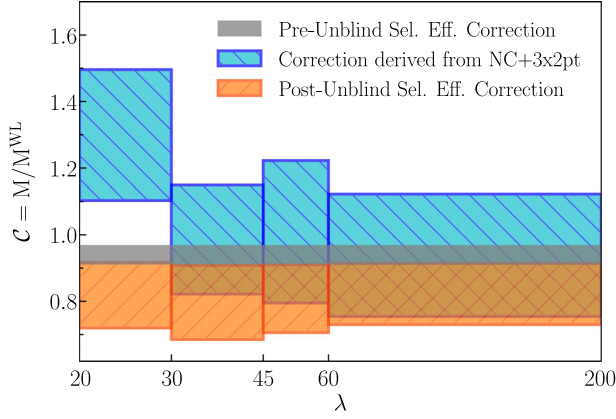


FIG. 12. Cyan bars: Mean correction required to reconcile the weak-lensing mass estimates from [15]—without the triaxiality and projection effects corrections—with the mean masses predicted by the combination of Y1 cluster counts and 3x2pt cosmology. Also overlotted the projection and triaxiality effects correction estimated analytically in [15] and adopted pre-unblinding (gray band), and the selection effect correction adopted post-unblinding (orange bars). The y extent of the bars represent the 68% confidence interval; the cyan bars are estimated as the ratio of the masses predicted by randomly sampling the NC + 3x2pt chain, and the “raw” weak-lensing masses randomly drawn from their posterior distribution.

weak lensing masses by  $1/(1 - f_{\text{cont}})$  and  $(1 - f_{\text{cont}})$ , respectively. Also, to account for a possible richness dependence we model the contamination fraction with a power law of the form:  $f_{\text{cont}}(\lambda^{\text{ob}}) = \Pi_0(\lambda^{\text{ob}}/25)^\pi$ . Finally, we fit for those parameters (along with all the others) combining cluster abundance and weak lensing data with DES 3x2pt cosmological priors, to derive the contamination fraction preferred by the our data sets in that cosmology. The fit results in a steeply decreasing contamination fraction ranging from  $\sim 15\%$  in the lowest richness bin to  $\sim 1\%$  in the highest richness bin. As expected, though, the model does not provide a good fit to the data in a 3x2pt cosmology, especially in the lowest richness bin where the predicted masses exceed the data by 15%–30%. Specifically, repeating the analysis without including the cosmological priors and fixing the contamination fraction parameters to their best-fit values, we obtain cosmological posteriors which are still at  $1.6\sigma$  tension with DES 3x2pt. Importantly, a high fraction of false detection at low richness and redshift is also disfavored by *Swift* x-ray follow up of  $\lambda \approx 30$  clusters, in which all but one of  $\approx 150$  low-richness ( $\lambda \in [25, 35]$  and  $0.1 < z < 0.35$ ) SDSS redMaPPer targets were x-ray detected (von der Linden *et al.*, in preparation).

One systematic that might seem like a good candidate for explaining the bias in Fig. 12 is the impact of baryonic processes: baryonic feedbacks redistribute and expel mass from a galaxy cluster, leading to cluster counts and weak-lensing masses that are biased low relative to expectations

from dark matter only simulations. Moreover, the effect would be stronger at low richness than at high richness, naturally producing a richness-dependent bias. However, results from hydrodynamical simulations disfavor this solution. If the triaxiality and projection effects are roughly mass independent, as found in Appendix D and per our *a priori* expectations, then the amplitude of the baryonic feedback would be  $\sim 30\%$  for clusters of richness  $\lambda \approx 25$ . That is, baryonic feedback would need to expel nearly 30% of the mass of a  $\sim 10^{14} M_\odot$  galaxy cluster, a fraction twice as large as its baryonic content ( $f_b \simeq \Omega_b/\Omega_m \simeq 0.15$ ), a clearly unphysical proposition [e.g., [60–63]]. Similarly, [56], using  $M > 10^{14} M_\odot$  clusters extracted from a hydrodynamical simulation, found that the redistribution of mass due to baryonic feedback processes induces a  $\sim 9\%$  bias on the recovered weak lensing mass, a factor of 3 times smaller than the bias required to reconcile our data sets in a 3x2pt cosmology. Moreover, we expect this bias to be further reduced in our analysis given that our fits allow the concentration parameter to vary with no informative priors in each bin, partially absorbing the effect of the mass redistribution.

Richness-dependent cluster miscentering suffers from much the same difficulty in explaining the observed discrepancy. While a systematic trend in cluster miscentering could introduce a richness-dependent bias in the recovered weak-lensing masses, it is hard to imagine miscentering giving rise to a 30% underestimate of the cluster mass. Such a correction would require a very high miscentering fraction at low richness, again in tension with *Swift* x-ray follow-up of low-richness SDSS redMaPPer clusters (von der Linden *et al.*, in preparation).

Cluster percolation has recently been identified as another possible source of systematic uncertainty [85]. Excessive percolation could give rise to severe incompleteness in the low-richness bins, as we found was needed to reconcile our final weak-lensing masses with the cluster counts within the context of a DES 3x2pt cosmology. If this were the case, then our percolation scheme must be overly aggressive. To test this, we reduce the percolation radius used from  $1.5R_\lambda$  to  $1.25R_\lambda$ . The corresponding change in the number of clusters is just under 1%, far from what would be needed to reconcile the cluster lensing and number counts data in a 3x2pt cosmology. We have also tested the impact of percolation on the weak lensing bias expected from numerical simulations, again finding a negligibly small impact.

In short, we have thus far been unable to identify a systematic that can plausibly explain the tension between the weak lensing data and the cluster counts assuming a DES 3x2pt cosmology (Fig. 10), particularly for our lowest-richness bins.

Interestingly, a lensing signal lower by  $\sim 20\%$ – $40\%$  compared to predictions from galaxy clustering has been measured by [86] around BOSS CMASS massive galaxies



at small scales ( $M \sim 10^{13} M_\odot$ ). If the discrepancy in their measurement were somehow related to the low weak lensing mass of our low richness clusters, that would point toward a mass-dependent physical origin for the bias that “turns on” around  $\approx 10^{14} M_\odot$ .

### C. Relation to other works

We have seen that the bias in the cosmological posterior shown in Fig. 6 can be fundamentally traced to the slope derived from our weak-lensing masses. Figure 15 of [15] compares the DES mass–richness relation to several other works in the literature. All of these tend to have relatively large slopes, though the DES value is unusually large. Two works in particular find slopes below unity: [27,87]. Of these, [87] has large error bars, so we will focus on the work by [81], which is an update to the [27] analysis.<sup>13</sup>

We use the method of Sec. V D to derive the mass–richness relation as constrained using cluster abundances when assuming a DES Y1 3x2pt cosmology. In Fig. 13 we compare this mass–richness relation (gray band) to that derived from our combined counts and weak-lensing analysis (cyan band), and to the mass–richness relation from [81] (magenta band). The latter is obtained as follows. First, they cross-match clusters selected using the Sunyaev–Zel’dovich (SZ) effect as measured using the South Pole Telescope (SPT) so that each SPT cluster is assigned a richness. Second, they assume a fiducial cosmology with  $\sigma_8 = 0.8$  and  $\Omega_m = 0.3$ . Using the SPT selection function, the abundance of clusters as a function of SZ-signal constrain the cluster masses, which in turn leads to a constraint of the richness–mass relation. In practice, this whole procedure is simultaneous and occurs at the likelihood level. It is worth noting that the SPT clusters typically have high richness values, with a median richness of 71. Thus, the constraint shown in Fig. 13 at low richness is an extrapolation of their results.

The agreement between the [81] analysis and the posterior obtained by analyzing the optical cluster abundance assuming a DES 3x2pt cosmology is remarkable. Given the similarity of the  $S_8$  values— $S_8 = 0.782 \pm 0.027$  for DES 3x2pt and  $S_8 = 0.8$  in the [81] analysis—this agreement implies that the optical and SPT abundances are compatible with each other, further strengthening the case that some unmodeled systematics reside with the interpretation of the stacked weak lensing signal as mean cluster mass rather than the modeling of the richness–mass relation. In particular, assuming a large incompleteness or contamination fraction as discussed above would result, for the combination of abundance data and DES 3x2pt cosmology

<sup>13</sup>The recent analysis of [88] also results in a shallower slope of the mass–richness relation, but their analysis includes assumptions about x-ray scaling relations and the scatter of the richness–mass relation, which make it more difficult to interpret their results within the context of our analysis.

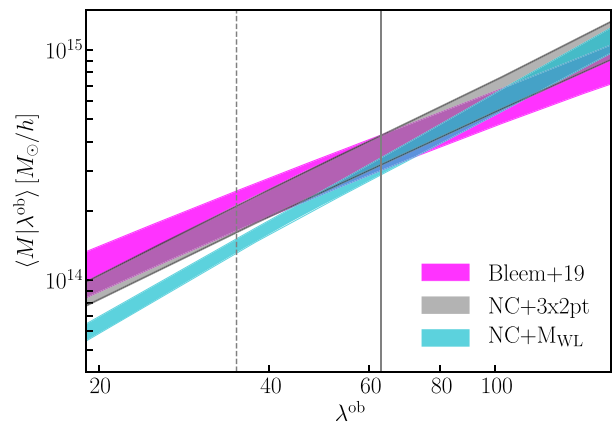


FIG. 13. Comparison of mass–richness relations at the mean sample redshift  $z = 0.45$ . The cyan and gray bands show the  $M$ – $\lambda^{\text{ob}}$  relation derived in this work combining number counts data with weak-lensing mass estimates or DES Y1 3x2pt cosmological prior, respectively. For comparison, shown in magenta the  $\langle M | \lambda^{\text{ob}} \rangle$  relation from SPT SZ clusters [81]. The dashed ( $\lambda^{\text{ob}} = 35$ ) and solid ( $\lambda^{\text{ob}} = 63$ ) vertical lines correspond to the richnesses above which 68% and 95% of the SPT-SZ sample used in [81] is contained.

priors, in a slope inconsistent with the results of [81]. Importantly, at  $\lambda \gtrsim 40$ —the richness range probed by the SPT sample—the weak-lensing masses and [81] results overlap. Consistent results are also obtained by Grandis *et al.* (in preparation), who use cross-matched redMaPPer–SPT clusters with  $\lambda > 40$  and the SZ signal–mass relation derived from the cosmological analysis of the SPT 2500 deg<sup>2</sup> cluster sample [9] to calibrate the richness–mass relation. Similarly to [81], when extrapolating their results to low richnesses ( $\lambda \lesssim 30$ ) the predicted cluster masses are  $\sim 30\%$  larger compared to our weak lensing mass estimates, while the predicted number counts are consistent with the redMaPPer abundance data.

Figure 13 is a modern incarnation of an old problem. Reference [89] studied the scaling relation between the richness of maxBCG clusters [90] and the SZ signature of those clusters using *Planck* data. They found both a large amplitude offset, and a large difference in the slope, relative to that predicted using weak-lensing masses. Reference [91] argued that the difference in amplitude was due primarily to the assumed *Planck* masses being biased low by  $\approx 30\%$ , and the weak-lensing masses being biased high by  $\approx 10\%$ . The difference in slope was, at that time, not significant given the corresponding uncertainties. This is related to the fact that, even though our analysis of the SDSS redMaPPer sample [19] undoubtedly suffers from the same systematics as our DES analysis, our SDSS results are consistent with the DES 3x2pt cosmology analysis. In other words, it is only because of the improved statistical constraining power of the DES that the “high” slope of the mass–richness relation derived using weak lensing is now clearly problematic.

To emphasize this point, we have rerun our analysis after dropping our lowest richness bin, making the mass range of our cluster sample more similar to that probed by x-ray and SZ selected catalogs. The resulting posteriors are shown in Fig. 14. As we can see, dropping our lowest richness bins

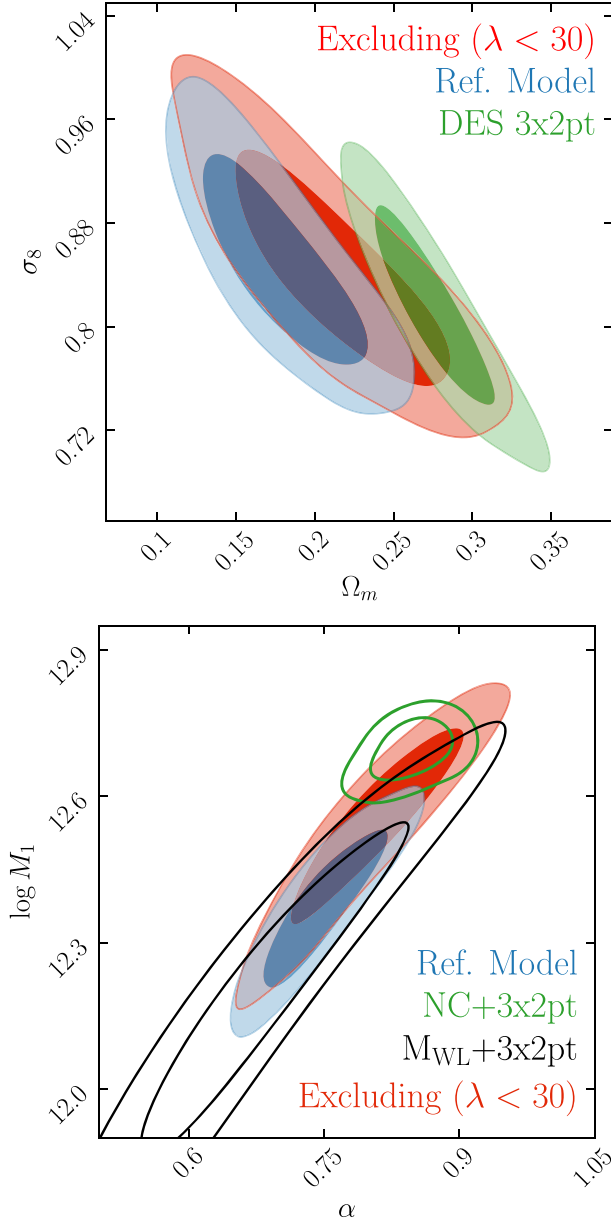


FIG. 14. Cosmological posteriors in the  $\sigma_8$ – $\Omega_m$  (upper panel) and  $\log M_1$ – $\alpha$  (lower panel) plane for our fiducial analysis (blue), and a new analysis in which we remove the lowest richness bins (red). Removing the low richness bins shifts the posteriors toward larger  $\Omega_m$  values, bringing our analysis into agreement with the DES Y1 3x2pt cosmology analysis ( $0.9\sigma$  tension; green contours in the upper panel). Similarly, when excluding the low richness bins, the richness-mass relation posteriors move toward the region of the parameter space preferred by the combination of DES number counts and 3x2pt cosmological priors (green contours in the lower panel).

shifts our posteriors toward higher matter density, bringing our analysis into agreement ( $0.9\sigma$ ) with the DES 3x2pt cosmology (upper panel). On the other hand, the posteriors of the richness-mass relation move toward the region of the parameter space preferred by the combination of number counts and DES 3x2pt priors (lower panel), and thus by the analysis of [81] using SZ selected clusters (see Fig. 13). Moreover, if we use the results of this analysis to predict our observables in the lowest richness bins, we obtain predictions for the number counts which are consistent with the abundance data, while the predicted mean cluster masses are higher by 15%–30% than the weak lensing mass estimates. These results highlight the fact that most of the tension with the DES 3x2pt cosmology is driven by the  $\lambda < 30$  data, and that our weak lensing mass estimates for  $\lambda < 30$  and  $\lambda > 30$  are inconsistent with each other within our model when combined with abundance data. Further removal of the next-lowest richness bin does not systematically shift the contours of the posterior. Aside from noting that our results are indeed especially sensitive to our lowest richness bin, Fig. 14 makes a simple but important point: had we performed our analysis with fewer, more massive clusters—analogously to previous abundance studies using X-ray and SZ selected clusters—the underlying systematic that biased the cosmological posteriors in Fig. 6 would have remained undiscovered. While this does not in any way demonstrate that clusters selected at other wavelengths will suffer from a similar systematic, it does open the possibility that such a systematic might exist also for low mass objects selected at other wavelengths.

One intriguing possibility that arises from this discussion is the extent to which the biases uncovered in our analysis could be mitigated using different mass-calibration strategies. For instance, in a recent work [88] used dynamical information to calibrate the richness-mass relation of galaxy clusters using the CODEX cluster sample. Encouragingly, they find a much shallower slope for the richness-mass relation, though their amplitude is in tension with ours and that of [81]. Of course, this does not negate the fact that the as-of-yet unidentified reason for discrepancy must be identified and understood, but it is encouraging to find that alternative methods of mass calibration may be less susceptible to the latter.

Another possibility resides in the use of different mass proxies. A stellar mass based mass proxy, such as the one presented in [92] is expected to be less impacted by projection effects [93]. In future work, we plan on comparing results from these different mass proxies, which could help with shedding light on the unknown systematics found in this work.

#### D. Correlated scatter

The analysis presented here is a “backward” analysis, in that one uses the weak lensing data to infer a cluster mass. This is to be contrasted to a “forward” analysis, in which one forward-models the weak-lensing shear profile of

galaxy clusters. Forward analyses [e.g., [7,9,10]] have traditionally assumed log-normal observable-mass relations, where the weak lensing signal is characterized by a weak-lensing mass  $M_{\text{WL}}$  that can correlate with the cluster-selection observable. In the presence of correlated scatter,  $P(M_{\text{WL}}|M, \lambda) \neq P(M_{\text{WL}}|M)$ . Instead, the expectation value of  $M_{\text{WL}}$  is still a log-normal distribution, but the mean is given by [94]

$$\langle \ln M_{\text{WL}} | \lambda \rangle = \langle \ln M | \lambda \rangle + \beta r \sigma_{M|\text{WL}} \sigma_{M|\lambda}, \quad (8)$$

where  $\beta$  is the slope of the halo mass function at the appropriate mass, and  $r$  is the correlation coefficient between the weak-lensing mass and the cluster observable.

Based on the above equations, it is easy to understand how the forward and backward modeling approaches are related. In the backward modeling approach, we consider the “correction term”  $\beta r \sigma_{M|\text{WL}} \sigma_{M|\lambda}$  to be an unknown for which we place priors based on numerical simulations. When  $r > 0$ , as expected from projection effects and triaxiality, this leads  $M_{\text{WL}}$  to be biased high.

There are two points to emphasize here. First, there is the simple equivalence of forward and backward modeling. A “forward model” with the same assumptions as we have would result in identical cosmological posteriors. Second, within the context of a log-normal model, Fig. 12 demonstrates that, under the assumption of the DES 3x2pt cosmology prior, the correlation coefficient between richness and weak-lensing mass must change as a function of mass, with  $r > 0$  at high mass (as expected), and  $r < 0$  at low mass. What can give rise to such a trend in the correlation coefficient remains unknown. Put another way, neither the “direction” of the analysis, nor the adoption of a multivariate log-normal model with correlated scatter, can resolve the tension in Fig. 6.

A second point of interest for forward modeling comes about because of the results shown in Appendix D. In particular, Fig. 17 demonstrates that selection biases may have strong scale dependence, and therefore cannot generally be modelled using a single “weak-lensing mass bias.” In a forward-model analysis, one should introduce the scale-dependent perturbations to the weak-lensing profiles, and marginalize over the amplitude of said perturbations. While we fully expect that an effective “weak-lensing mass” suffices for now, we expect future cluster analyses will require an understanding of the scale-dependent impact of selection effects (or, within the context of a log-normal model, an understanding of the scale dependence of the weak-lensing scatter and correlation coefficient).

## VII. SUMMARY AND CONCLUSION

We have performed a cosmological analysis of the abundance and the weak lensing signal of the DES Y1 redMaPPer clusters. We summarize our findings below:

- (i) The cosmological posteriors of our unblinded analysis are in  $5.6\sigma$  tension with *Planck* CMB, and  $2.4\sigma$  tension with the DES 3x2pt cosmological analysis in the  $\sigma_8$ - $\Omega_m$  plane. This is driven by a low  $\Omega_m$  posterior that is in tension with all existing cosmological probes. This finding is robust to the adopted cosmological and richness-mass relation model.
- (ii) The internal inconsistency of the DES Y1 cluster data with other DES probes rule out the possibility that the tension is driven by an observational systematic affecting the DES data.
- (iii) Cross checks of the redMaPPer catalog with X-ray and SZ data suggest that the abundance data and related modeling are not driving the tension but it is likely a consequence of an incorrect interpretation of the stacked weak lensing signal of the DES redMaPPer clusters.
- (iv) Low richness data ( $\lambda \in [20, 30]$ ) are the main driver of the tension with the DES 3x2pt cosmological results: dropping our lowest richness bin from the analysis removes the tension with DES 3x2pt ( $0.9\sigma$ ). In particular, the weak lensing mass estimates for  $\lambda < 30$  push the slope and amplitude posteriors of the richness-mass relation toward lower values compared to the ones preferred by the combination of number counts and weak lensing data at  $\lambda > 30$ , as well as by the analysis of [81] using SPT clusters ( $\lambda \gtrsim 40$ ).
- (v) Assuming our abundance data, modeling and DES 3x2pt results to be correct, we estimate the required bias in the observed weak-lensing masses by comparing the latter to the predicted masses assuming a DES 3x2pt cosmology and using the cluster counts to constrain the richness-mass relation. The relative mass offset we recover is richness dependent, corresponding to a steeper slope in the richness-mass relation compared to the one preferred by the weak lensing data.
- (vi) Our understanding of how photometric cluster selection impacts the stacked lensing profiles of clusters might have a major role in the observed tension. However, at low richness, the necessary selection effect bias requires the raw weak-lensing masses of photometrically selected clusters to be biased low relative to a mass-selected sample. This is contrary to our *a priori* expectations, and we have not yet been able to identify a systematic that could give rise to such a selection effect.
- (vii) Interpreting our results within the context of correlated observables, our data implies that the correlation coefficient between richness and weak lensing is mass dependent, and changes sign in going from high mass clusters (positive correlation) to low mass clusters (negative correlation). As noted above, this is very surprising.



As discussed in Sec. VIC, hints of a richness-dependent bias in the weak lensing signal of galaxy clusters go as far back as [89], but it is only with the improved statistical power of the DES that these biases have become statistically significant. Understanding the origin of this systematic effect, and the degree to which it can be calibrated using multi-wavelength cluster data, is an absolute necessity for future photometric cluster cosmology analyses. Observational and simulation-based campaigns to study the relation of true cluster mass, observed richness, and weak lensing profiles, independent of the inherent limitations of purely photometric data, will shed light on the puzzles posed by DES Y1 cluster abundance and lensing data.

### ACKNOWLEDGMENTS

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. 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. The DES data management system is supported by the National Science Foundation under Grants No. AST-1138766 and No. AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under Grants No. AYA2015-71825, No. ESP2015-66861, No. FPA2015-68048, No. SEV-2016-0588, No. SEV-2016-0597, and No. MDM-2015-0509, some of which include ERDF funds from the European Union. I. F. A. E. 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 No. 240672, No. 291329, and No. 306478. We acknowledge support from the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through Project No. 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. M. C. and A. S. are supported by the ERC-StG “ClustersXCosmo” Grant agreement No. 716762. A. S. is supported by the FARE-MIUR grant “ClustersXEucld”. E. R. was supported by the DOE Grant No. DE-SC0015975, by the Sloan Foundation, Grant No. FG-2016-6443, and the Cottrell Scholar program of the Research Corporation for Science Advancement. This research used simulations that were performed resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

### APPENDIX A: CALIBRATION OF THE DISTRIBUTION $P(\lambda^{\text{ob}}|\lambda^{\text{true}}, z^{\text{true}})$

A key ingredient in our analysis is our characterization of noise in photometric richness estimates. As discussed in Sec. IV, we consider three distinct sources of noise in  $\lambda^{\text{ob}}$ :

- (1) A Gaussian random noise associated with photometric uncertainties, uncorrelated structures, and background subtraction.
- (2) An exponentially decaying additive contribution to the richness due to projection effects that is dominated by the contribution from correlated structures along the line-of-sight.



- (3) A multiplicative correction that removes galaxies from the cluster richness estimates of a small fraction of low mass systems due to the impact of percolation in the construction of the redMaPPer cluster catalog.

The random noise can straightforwardly be calibrated from the data. We use the matched filter used to detect redMaPPer clusters to generate Monte Carlo realizations of our cluster model, and proceed to insert these artificial clusters into our data set. We generate  $10^4$  cluster realizations along a grid of cluster richness  $\lambda^{\text{true}}$  and cluster redshift  $z$ . Each of these clusters is placed at a random point within the survey footprint, and the magnitude of every galaxy in the simulated cluster is perturbed according to the effective survey depth in each band at the galaxy's location. We then estimate the richness of the galaxy clusters. The distribution  $P(\lambda^{\text{ob}}|\lambda^{\text{true}})$  obtained in this way is very well approximated as a Gaussian, and the observational uncertainty on the posteriors of these parameters is negligible. In this way, we fully characterize observational uncertainties due to photometric uncertainties, uncorrelated structures, and background subtraction.

We characterize the impact of correlated large scale structure using the method developed in [36]. This model is intuitively very simple: when two clusters are aligned along the line of sight, the smaller of the clusters will get blended into the larger of the two systems, with some fraction of its galaxies being mistakenly assigned to the larger system. The fraction of galaxies that the small cluster loses will depend on the separation along the line of sight between the two systems: if the separation is zero, the smaller cluster will be entirely subsumed within the larger system, while if the separation is large the two clusters will be easily distinguished from each other, so there will be no artificial projection effects. Evidently, the critical input to this model is the calibration of how the strength of projection effects decreases with increasing cluster separation. Note that the fraction of the cluster lost to projection effects must be unity at zero separation, zero at large separation, and must have a slope of zero at zero separation. Consequently, we expect *a priori* that a simple Gaussian can successfully describe this function.

We calibrate the separation dependence of projection effects by calculating the fractional decay of the cluster richness as a function of redshift, that is, the fraction of member galaxies of a cluster that redMaPPer would assign to a putative cluster perfectly aligned with the former as a function of their separation in redshift. This fractional decay is in fact well described by a Gaussian, enabling us to calibrate the width of this Gaussian as a function of cluster redshift. Because this function should reflect only the intrinsic width of the red sequence and photometric errors, we did not expect this fraction to depend on cluster richness, an expectation that we explicitly confirmed. We then calibrated the width of the Gaussian decay as a

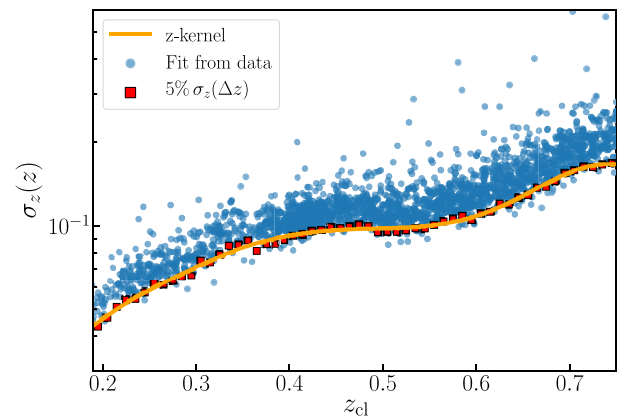


FIG. 15. The blue dots are the best-fit values for  $\sigma_z$  obtained when fitting the curves  $\lambda(z)$  for each cluster in the redMaPPer cluster catalogue. The red squares represent the 5 percentile of the  $\sigma_z$  distribution estimated in redshift bins of width  $\Delta z = 0.01$ . The solid orange line shows the model for  $\sigma_z(z)$  adopted for the analysis.

function of redshift in the DES data. The resulting calibrated data is shown in Fig. 15. Our best-fit model is a simple polynomial fit that successfully described our data.

We use our projection effects model to generate a synthetic data set as follows. Starting from the DES Buzzard light cone simulation (DeRose *et al.* 2018, in prep, Wechsler *et al.* 2018, in prep), we assign to each halo an intrinsic richness  $\lambda^{\text{true}}$ . We then rank order the halos by  $\lambda^{\text{true}}$ , and proceed to compute their projected richnesses using the projection effect model of [36] as calibrated above. Halos that contribute a fraction  $f$  of their galaxies to a richer system along the line-of-sight have their own final richness decreased by a factor  $1 - f$ , i.e., we enforce galaxy conservation. The end result is a galaxy cluster catalog that includes both projection effects and percolations. We use this simulated catalog to characterize both the characteristic richness enhancement due to projection effects which characterizes the exponential distribution of this noise, and the fraction of galaxy clusters that suffer from percolation effects (i.e., the fraction of clusters who lost some of their galaxies to richer systems along the line of sight). Both of these effects are richness and redshift dependent: richer systems live in denser environments, which increases the importance of projection effects. Likewise, systems at higher redshift are noisier, making it easier to blend systems together, and therefore increasing the impact of projection effects. Finally, with regards to percolation, low richness systems are much more likely to suffer from percolation effects (the richest systems rarely have an even richer system along their line of sight).

These trends are all very precisely measured in the simulation, and the corresponding observational uncertainties are negligible compared to the associated systematic uncertainties. In particular, it should be obvious that the impact of projection effects is cosmology dependent:

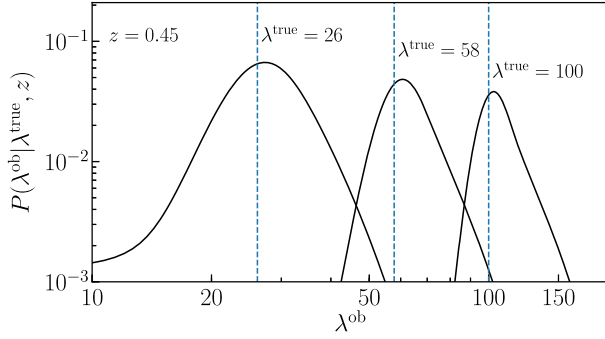


FIG. 16.  $P(\lambda^{\text{ob}}|\lambda^{\text{true}})$  distribution for clusters of true richness  $\lambda^{\text{true}} = 26, 58$ , and  $100$  at the mean redshift of the sample  $z = 0.45$ .

higher  $\sigma_8$  and higher  $\Omega_m$  models will result in increased projection effects. Fortunately, as demonstrated in [36], these differences are very nearly degenerate with the parameters of the intrinsic richness-mass relation, so the cosmological posteriors from our analysis are extremely robust to these types of effects. Indeed, as we demonstrate in the main body of this text, even if we entirely neglect the impact of correlated structures along the line of sight, our cosmological posteriors are hardly affected.

Figure 16 shows our calibration of the distribution  $P(\lambda^{\text{ob}}|\lambda^{\text{true}})$  for clusters of richness  $\lambda^{\text{true}} = 20, 58$ , and  $100$  at the mean redshift of the sample  $z = 0.45$ . The Gaussian peak due to observational noise is evident, as is the long-tail to high richness due to projection effects. The low tail at low richness is due to percolation.

## APPENDIX B: BLINDING AND UNBLINDING PROTOCOL

The DES analysis was performed blind, but followed a staged unblinding procedure. In particular, we used the DES inference pipeline to analyze the SDSS redMaPPer cluster catalog [19]. Both the SDSS and DES analyses were performed blind, but the unblinding of the SDSS analyses was part of the DES unblinding protocol, as described below. This staged unblinding has both benefits and drawbacks. The principal benefit is that “unknown unknowns” may be discovered and fixed after unblinding a “weak” data set (SDSS), enabling us to implement any necessary corrections before unblinding the “strong” (DES) data set. The principal drawback is that this type of blinding is somewhat weaker than simply unblinding the “strong” data set. However, we emphasize that the DES photometry, shear, and photo- $z$  catalogs are completely independent of the corresponding SDSS catalog. In practice, no “unknown unknowns” were identified when unblinding the SDSS data sets, so the effective path through the DES unblinding protocol was identical to that of a simultaneous unblinding.

Our DES blinding protocol is as follows:

- (1) The cosmological parameters in the MCMC were randomly displaced before being stored. The displacement was stored in a not-human-readable format (binary).
- (2) All modeling choices for both SDSS and DES were made before unblinding of the SDSS data sets. Modeling choices were not allowed to change after unblinding of the SDSS data set.
- (3) In addition to our fiducial model for the scaling relation, we considered one additional model for projection effects, namely random-point-injection. Random-point-injection refers to the projection effects model calibrated by inserting galaxy clusters at random locations in the sky. This method obviously underestimates the impact of projection effects, so we take half of the difference in cosmological parameters between our fiducial model and this extreme projection effects model as the systematic uncertainty associated with modeling of projection effects.
- (4) All priors for both the SDSS and DES data sets were finalized before unblinding of the SDSS data set, with one critical exception: the prior on the intrinsic scatter parameter  $\sigma_{\text{intr}}$ . In [19], we applied a prior  $\sigma_{\text{intr}} \in [0.1, 0.8]$ . At  $\sigma_{\text{intr}} = 0.8$ , the model predicts that  $\approx 11\%$  of massive clusters ( $M \geq 10^{15} M_\odot$ ) do not host red-sequence galaxies ( $\lambda^{\text{true}} = 0$ ). This seems implausible. Reference [95] studied the scatter in richness at fixed mass in numerical simulations in which galaxies were used to populate resolved halo substructures. They then fit Gaussian distributions to their results. Their best fit total fractional scatter in a maxBCG-like cluster catalog [90] was  $0.37$ .<sup>14</sup> Note this is a total scatter, so  $\sigma_{\text{intr}}$  must be strictly less than  $0.37$  in this simulation. Moreover, since redMaPPer is demonstrably superior to maxBCG [24], the above number should be pessimistic. Based on this argument, we set for the DESY1 analysis the conservative upper limit  $\sigma_{\text{intr}} \leq 0.5$ . This upper limit is also low enough that fluctuations that produce negative richnesses (formally nondetections) are sufficiently rare for them to be irrelevant for our study ( $P(\lambda^{\text{true}} \leq 0) \leq 0.6\%$ ).<sup>15</sup>

Finally, the DES analysis includes an additional parameter,  $\epsilon$ , governing the redshift evolution of the intrinsic richness-mass relation (see Equation (7)).

<sup>14</sup>Tails due to projection effects were obvious, but we note that our model explicitly accounts for such tails.

<sup>15</sup>A Gaussian model is mathematically more convenient than a log-normal model, both because Poisson distributions are closer to a Gaussian distribution than to a log-normal distribution, and because the exponential model for projection effects is easy to convolve with a Gaussian.

- (5) The weak-lensing masses of the clusters in each richness/redshift bin remained blind throughout the entire weak lensing analysis, which was completed before SDSS unblinding. No alterations of the lensing pipeline were made post-unblinding of the DES weak lensing data, except for a minor bug-fix that affected the boost factor correction of only one richness/redshift bin. The change in mass was well below the uncertainty for that bin, and the bug was found and fixed before unblinding the cosmological constraints. For details on our weak lensing calibration of the DES data set, we refer the reader to [15].
- (6) No comparison of our cosmological constraints to any other data sets were performed prior to unblinding of the DES data. Our analysis in [15] did compare the DES and SDSS weak-lensing masses, but this was only done after the DES masses were unblinded.

Our unblinding protocol was defined by the set of requirements detailed below.

- (1) The SDSS analysis was unblinded, and “unknowns unknowns” were either not found or addressed, as appropriate.
- (2) All noncosmological systematics tests of the shear measurements were passed, as described in [37], and all priors on the multiplicative shear biases were finalized.
- (3) Photo- $z$  catalogs were finalized and passed internal tests, as described in [48].
- (4) Our inference pipeline successfully recovered the input cosmology in a synthetic data set (see Appendix E).
- (5) All planned DES-only chains (including alternative models) were run and satisfied the Gelman-Rubin convergence criteria [96] with  $R - 1 \leq 0.03$ .
- (6) Since not explicitly included in the analysis, we demanded that the systematic uncertainty in our posteriors due to projection effects modeling—estimated as half the difference between the central values of the posteriors for our fiducial model and the random-point-injection model—were smaller than the corresponding statistical uncertainties.
- (7) We verified that the posteriors of all parameters which we expected would be well-constrained did not run into the priors within the 95% confidence region when using a flat  $\Lambda$ CDM model. Parameters that are prior dominated are  $M_{\min}$ ,  $\sigma_{\text{intr}}$ ,  $s$ ,  $q$ ,  $h$ ,  $\Omega_b h^2$ ,  $\Omega_c h^2$ , and  $n_s$ . All of these were expected to be prior dominated, and all prior ranges were purposely conservative. Of these, the two that might be most surprising to the reader might be  $M_{\min}$  and  $\sigma_{\text{intr}}$ , as these parameters help govern the richness-mass relation. However, notice that  $M_{\min}$  is the mass at which halos begin to host a single central galaxy; since our cluster sample is defined with the richness

threshold  $\lambda \geq 20$ , the mass regime of halos which host a single galaxy is simply not probed by our data set. Likewise, our data vector is comprised only of the mean mass of galaxy clusters in a given richness bin, a quantity that is largely independent of the scatter in the richness-mass relation [see [15], which accounts for the modest scatter dependence as a systematic uncertainty in the recovered masses].<sup>16</sup>

- (8) Finally, this paper underwent internal review by the collaboration prior to unblinding. All members of the DES cluster working group, as well as our internal reviewers, had to agree that our analysis was ready to unblind before we proceeded.

## APPENDIX C: BLINDED ANALYSIS RESULTS

After all of our unblinding requirements were satisfied, we proceeded to unblind our results. For the two cosmological parameters constrained by our data set we obtained for the *blinded* analysis  $\Omega_m = 0.172^{+0.023}_{-0.029}$  and  $\sigma_8 = 0.956^{+0.045}_{-0.056}$ , corresponding to  $S_8 = 0.720 \pm 0.032$ . Figure 19 shows the resulting posteriors on the  $\sigma_8$ - $\Omega_m$  plane (*blinded* analysis; gray), along with the posteriors obtained from the *unblinded* analysis (i.e., our reference results; red), DES 3x2pt (blue) and *Planck* CMB (gold). As can be seen from the figure, the *blinded* analysis results are in clear tension with those derived by the other DES probes and *Planck* CMB ( $2.3\sigma$  and  $6.7\sigma$  in the  $\sigma_8$ - $\Omega_m$  plane, respectively). Driven by the low  $\Omega_m$  value recovered, a larger than  $3.5\sigma$  tension is also present with BAO measurements [65–67] and supernovae data [68].

The  $\chi^2$  of the best-fit model of the *blinded* analysis is 38.35. Based on the expected  $\chi^2$  distribution (see Sec. VA for details) the model adopted in the *blinded* analysis did not provide a good fit to the data ( $\chi^2/\nu_{\text{eff}} = 38.35/18.65$ ). This was driven primarily by the offset between the predicted and observed abundances of galaxy clusters in our highest redshift and largest richness bin.

Given the large tension with the DES 3x2pt and *Planck* results, as well as with BAO, supernovae and other independent cluster count analyses, we attempted to trace back the source of the tension, whether it be an objective bug in the code and/or an unknown/underestimated source of systematic bias.

## APPENDIX D: POST-UNBLINDING TESTS AND SELECTION EFFECT CALIBRATION

Two minor bugs were discovered in our pipeline post-unblinding. First, the projection effect correction adopted in [15] was implemented with the wrong sign, and second we implement in our pipeline  $\langle \ln M | \lambda \rangle$  rather than  $\ln \langle M | \lambda \rangle$ .

<sup>16</sup>Interestingly, in the log-normal model the data does constrain the scatter parameter.



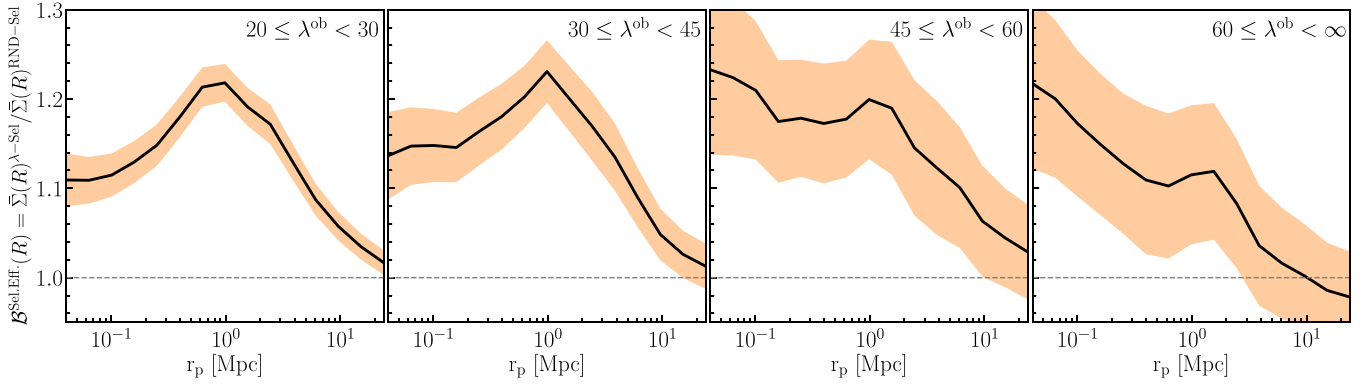


FIG. 17. Selection effect bias on the stacked mass density profile derived from synthetic clusters in the redshift range  $0.35 < z < 0.50$  for the four  $\lambda^{\text{ob}}$  bins considered in the analysis. The bias is computed as the ratio of the stacked  $\bar{\Sigma}(R)$  profiles measured in clusters selected by richness and clusters randomly selected from the simulations so as to match the mass and redshift distribution of the  $\lambda^{\text{ob}}$ -selected sample. The black lines correspond to the means of the biases retrieved from 12 Y1-like simulations, while the shaded area represent one standard deviation of the mean. The correction is included in to the stacked weak lensing model of [15] by multiplying the projected mass density profile  $\bar{\Sigma}(R)$  by the selection effect bias profile  $\mathcal{B}^{\text{Sel.Eff.}}(R)$  relevant for each bin prior to computing  $\Delta\Sigma(R)$ .

Fixing these bugs had only a minor impact on the cosmological posteriors. Post unblinding, an independently coded version of our cosmological pipeline was completed. The two pipeline codes were found to be in excellent agreement with each other, precluding the possibility of a bug in the code used to analyze the data (the bugs above came from the processing of the data).

To address possible model systematic biases we test a variety of different models for  $P(\lambda^{\text{ob}}|M, z)$  besides those considered pre-unblinding, which include:

- (1) A model in which the intrinsic scatter of the richness-mass relation is allowed to be mass dependent.
- (2) A model in which the intrinsic scatter of the richness-mass relation is allowed to be redshift dependent.
- (3) A model in which the slope of the mass-richness relation  $\alpha$  in Eq. (7) depends on mass:  $\alpha(M) = \alpha_0 + \alpha_M \log(M/10^{14.2})$ .
- (4) A model in which the slope of the mass-richness relation  $\alpha$  can evolve with redshift:  $\alpha(z) = \alpha_0[(1+z)/(1+z_*)]^{\alpha_z}$ .

None of the models tested seem to suggest a large systematic bias on cosmological posteriors related to model assumptions: the differences between  $P(\lambda^{\text{ob}}|M)$  models are mainly accommodated by a shift of the richness-mass relation parameters.

Finally, as noted in the main text, we used numerical simulations to update our model for the impact of selection effects on the recovered weak-lensing mass of galaxy clusters. This work was started before unblinding, but was only completed post-unblinding and found an effect in excess of previous literature results. We describe our calibration of selection effects below.

We ran redMaPPer on 12 simulated Y1-like light-cones from the Buzzard Flock suite [64]. The synthetic data have been tuned to match the observed evolution of galaxy

counts at different luminosities as well as the spatial clustering of the galaxy population of DES Y1 data. To avoid double counting of miscentering effects, redMaPPer has been run fixing the cluster center on the dark matter halo center. We thus computed the azimuthally averaged stacked mass density profile  $\bar{\Sigma}(R)$  of the clusters in richness/redshift bins using the dark matter particle distribution. Then, for each richness and redshift bin, we randomly selected 1000 halos from the simulations with the same mass and redshift distribution as the clusters in the bin. Finally, we measured the stacked mass density profiles of mass-selected cluster samples and compare them to those obtained from the richness/redshift selected samples.

Figure 17 shows the ratio of stacked mass density profiles of the redMaPPer selected clusters to that of the mass-selected sample. Note that the mass and redshift distribution of the two samples is identical by construction, thus any difference between the two is due to selection effects. We find that redMaPPer selected clusters have a weak lensing signal that is biased higher by  $\approx 10\text{--}25\%$  over the relevant radius range than that of similar, purely mass/redshift selected clusters<sup>17</sup> (see Fig. 17). This indicates that, at a given halo mass, redMaPPer preferentially selects halos with a boosted lensing signal compared to a random sample. This bias is partially due to triaxiality and projection effects, and will be studied in greater detail in an upcoming publication (Wu *et al.*, in preparation). Specifically, matching our control samples not only by mass and redshift, but also by the  $\sigma_z$  (our proxy for projection effects; see Appendix A) and halo orientation distributions of the richness-selected sample, reduces the bias between the two samples by  $\sim 50\text{--}100\%$  depending on the bin and radius considered.

<sup>17</sup>During the finalization of this analysis similar findings have been presented in the work of [97]



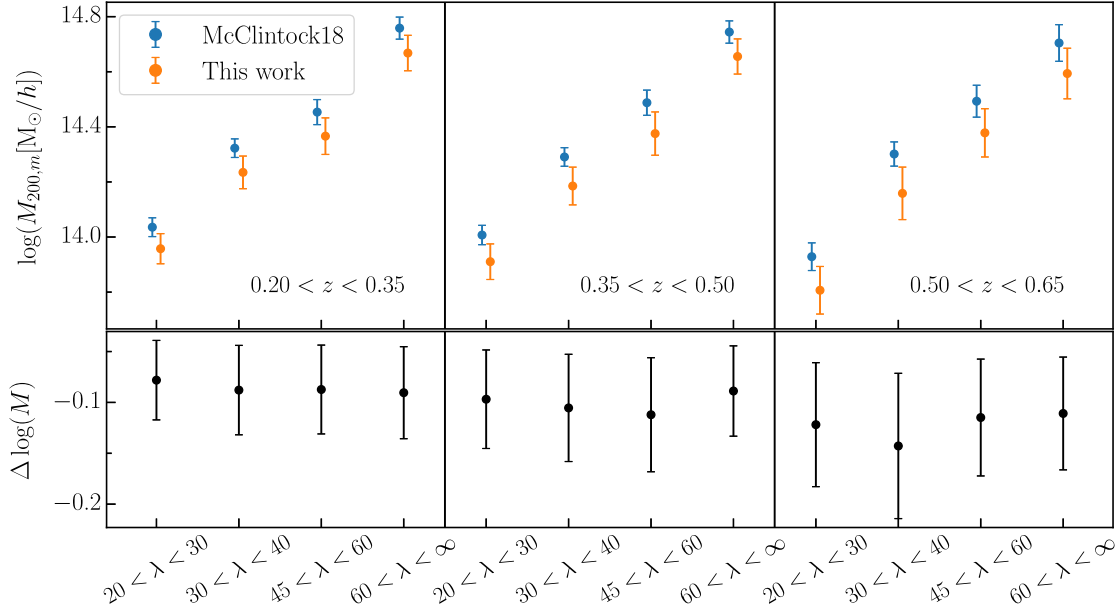


FIG. 18. Upper panel: Comparison of the weak-lensing masses derived in [15] and the ones adopted in this work which include the selection effect bias correction and uncertainty. The inclusion of this systematic lowers the weak-lensing mass estimates by 20%–30% and increases the error budget by 50%–60% depending on the richness/redshift bin. Lower panel: Difference of the log masses derived including or not the selection effect bias correction. The error bars correspond to the uncertainty associated with the selection effect bias estimated for each bin as half of the difference between the two mass estimates.

To account for this effect in our mass estimates we refit the stacked weak lensing data following the same methodology detailed in [15], but multiplying the projected mass density profile model  $[\tilde{\Sigma}(R)$ , Eq. 28 of [15]] by the selection effect bias correction relevant for the bin considered:  $\mathcal{B}^{\text{Sel.Eff.}}(R) = \tilde{\Sigma}(R)^{\lambda-\text{Sel}} / \tilde{\Sigma}(R)^{\text{RND-Sel}}$  prior to computing the predicted  $\Delta\Sigma(R)$ . The masses derived including this systematic are  $\sim 20\text{--}30\%$  smaller compared to the previous results, with a larger bias for richer and high redshift clusters (see Fig. 18). Since this systematic effect is still under investigation we add in quadrature to the refitted masses a conservative error equal to half of the difference between the old and the new results— $\sigma^{\text{Sel.Eff.}} = |M^{\text{new}} - M^{\text{old}}|/2$ —that is, the absence or doubling of selection effects on weak lensing mass estimates is excluded at  $2\sigma$ . While the mock redMaPPer catalogs obtained from the Buzzard Flock suite are known to underestimate the richness of the clusters at fixed mass [64], since the selection effect correction is calibrated on the relative bias of stacked lensing profiles of samples with the same mass distribution, we do not expect this to affect our results. As a confirmation of the latter statement, we repeat the analysis on synthetic redMaPPer clusters extracted from a Buzzard simulation adopting different assumptions for the red-sequence and clustering model, finding results consistent with the one above. Nonetheless, additional tests on different synthetic data will be fundamental to further validate our findings and reduce the associated uncertainty.

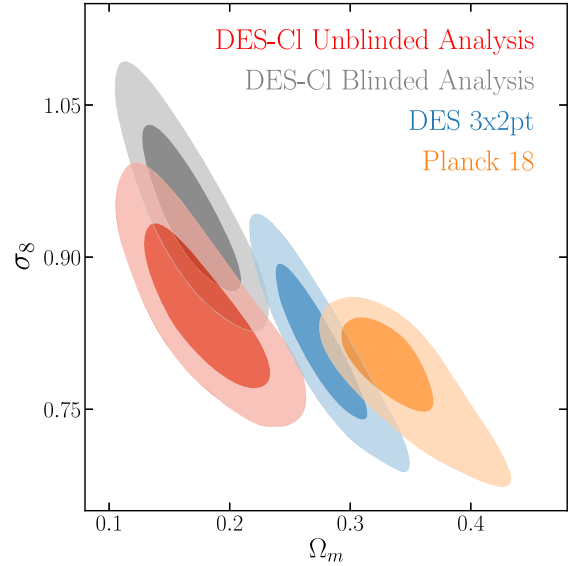


FIG. 19.  $\sigma_8$ – $\Omega_m$  posteriors from the DES Y1 cluster blinded analysis (gray) and unblinded analysis (red). The latter adopt the selection effect bias correction on the weak lensing mass estimates detailed in Appendix D. Shown for comparison also are the posteriors obtained from the DES 3x2pt (blue) and *Planck* CMB (gold) analysis. The smaller weak lensing masses recovered including the selection effect bias lead to a  $\sim 2\sigma$  shift of the  $\sigma_8$  posterior, while the larger systematic error associated to the masses entails a relaxation of the  $S_8$  posterior of  $\sim 18\%$ . As evident from the figure, the inclusion of the selection effect bias does not substantially change the level of tension with DES 3x2pt or *Planck* CMB results in the  $\sigma_8$ – $\Omega_m$  plane.

The cosmological constraints in the main text adopt this systematic calibration, as noted in Sec. III C.

In Fig. 19 we show the effects of the selection effect bias and associated systematic uncertainty on the DES Y1 cluster posteriors in the  $\sigma_8$ - $\Omega_m$  plane. The  $\sim 20\%$  lower weak-lensing mass estimates adopted in our reference analysis shift downwards by  $\sim 2\sigma$  the  $\sigma_8$  posteriors while leaving the  $\Omega_m$  posterior mostly unaffected (compare gray and red contours). Furthermore, the larger systematic error associated with the mass estimates causes the  $\sigma_8$  posteriors to relax by  $\sim 18\%$ . We note that the inclusion of the

selection effect bias does not substantially affect the level of consistency of our results with DES 3x2pt or *Planck* CMB posteriors. We further stress that the gray contours are shown here only to illustrate the effect of the selection bias, and should not be considered as possible alternative results of this analysis.

## APPENDIX E: PIPELINE VALIDATION

We validate our analysis pipeline using the simulated cluster catalog described in Appendix A. Specifically,

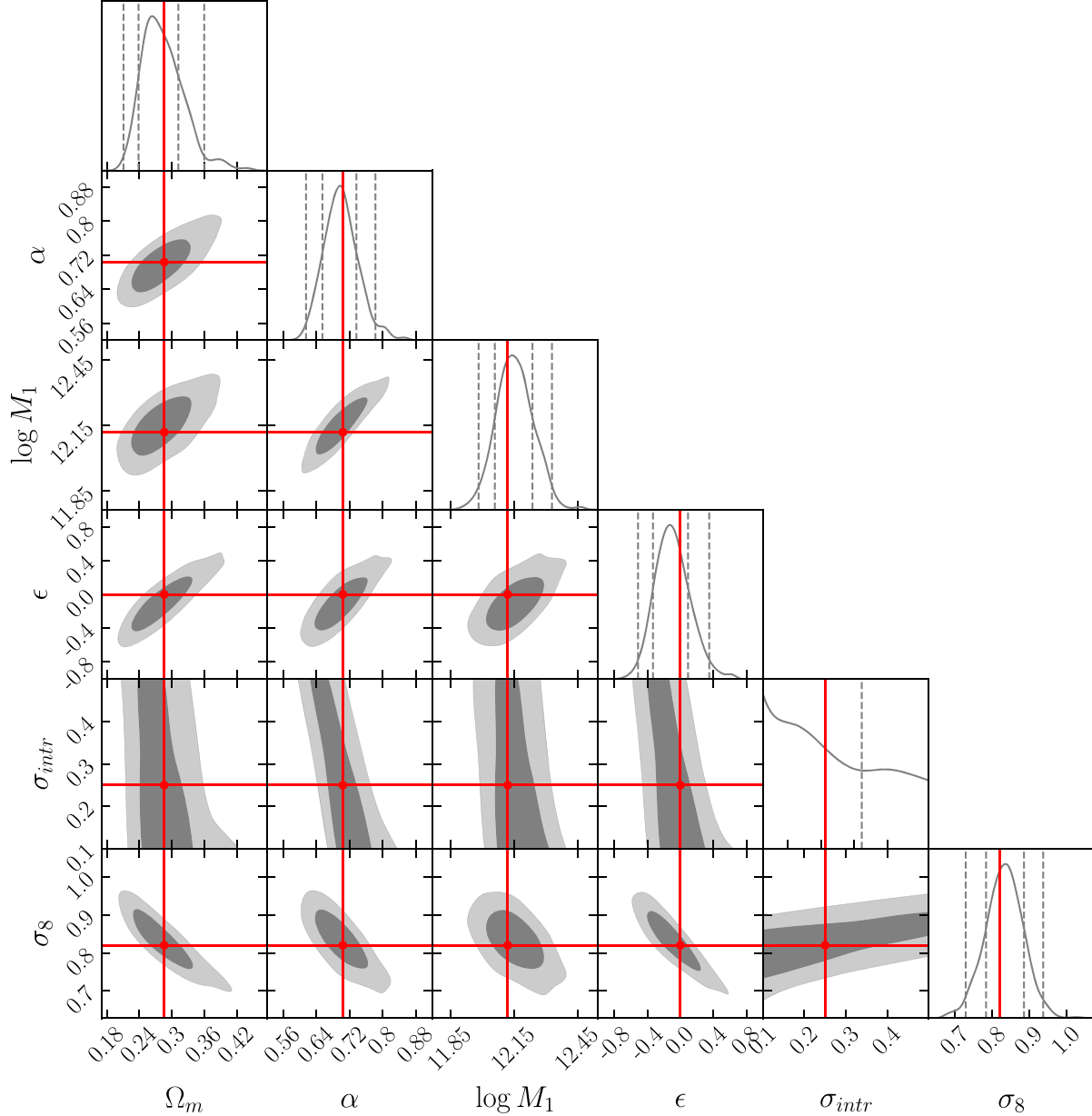


FIG. 20. 68% and 95% confidence contours obtained running our pipeline on mock data. The input parameter values used to generate the simulation and the mock data catalog are shown in red. The dashed lines shown in the 1-d marginalized distributions (diagonal of the triangle plot) correspond to the 0.025, 0.16, 0.84 and 0.975 quantiles of the distributions. This plot includes only the model parameters that are not prior dominated.

starting from the simulated cluster catalog described above, we bin the simulated clusters in richness and redshift bins in a way that is identical to that done in the real data. We then calculate the mean halo mass of the resulting galaxy clusters, and scatter it according to the observational noise in our cluster mass calibration. This “noise-scattering” properly accounts for correlated uncertainties due to systematics. The end result is a simulated data vector of cluster abundances

and mean cluster masses that can be used as an input to our cosmology pipeline. Figure 20 shows the posteriors from our pipeline when applied to this simulated data set. The input cosmology is shown as the intersection of the horizontal and vertical lines in each plane, which describe the input parameters of the simulation. The good agreement between these input parameters and our analysis posteriors demonstrate that our pipeline is working as intended.

- 
- [1] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolte, M. Halpern, R. S. Hill, N. Odegard *et al.*, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological parameter results, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013).
  - [2] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, [arXiv:1807.06209](https://arxiv.org/abs/1807.06209).
  - [3] B. Flaugher, H. T. Diehl, K. Honscheid *et al.*, The dark energy camera, *Astron. J.* **150**, 150 (2015).
  - [4] S. W. Allen, A. E. Evrard, and A. B. Mantz, Cosmological parameters from observations of Galaxy clusters, *Annu. Rev. Astron. Astrophys.* **49**, 409 (2011).
  - [5] A. V. Kravtsov and S. Borgani, Formation of Galaxy clusters, *Annu. Rev. Astron. Astrophys.* **50**, 353 (2012).
  - [6] R. A. Burenin and A. A. Vikhlinin, Cosmological parameters constraints from galaxy cluster mass function measurements in combination with other cosmological data, *Astron. Lett.* **38**, 347 (2012).
  - [7] A. B. Mantz, A. von der Linden, S. W. Allen, D. E. Applegate, P. L. Kelly, R. G. Morris, D. A. Rapetti, R. W. Schmidt *et al.*, Weighing the giants—IV. Cosmology and neutrino mass, *Mon. Not. R. Astron. Soc.* **446**, 2205 (2015).
  - [8] M. Cataneo, D. Rapetti, F. Schmidt, A. B. Mantz, S. W. Allen, D. E. Applegate, P. L. Kelly, A. von der Linden, and R. G. Morris, New constraints on  $f(R)$  gravity from clusters of galaxies, *Phys. Rev. D* **92**, 044009 (2015).
  - [9] S. Bocquet *et al.* (SPT Collaboration), Cluster cosmology constraints from the 2500 deg<sup>2</sup> SPT-SZ survey: Inclusion of weak gravitational lensing data from Magellan and the Hubble space telescope, *Astrophys. J.* **878**, 55 (2019).
  - [10] R. Murata, T. Nishimichi, M. Takada, H. Miyatake, M. Shirasaki, S. More, R. Takahashi, and K. Osato, Constraints on the mass-richness relation from the abundance and weak lensing of SDSS clusters, *Astrophys. J.* **854**, 120 (2018).
  - [11] C. Hikage, M. Oguri, T. Hamana, S. More, R. Mandelbaum, M. Takada, F. Köhlinger, H. Miyatake, A. J. Nishizawa *et al.*, Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data, *Publ. Astron. Soc. Jpn.* **71**, 43 (2019).
  - [12] P. Melchior, D. Gruen, T. McClintock, T. N. Varga, E. Sheldon, E. Rozo, A. Amara, M. R. Becker, B. A. Benson, A. Bermeo *et al.*, Weak-lensing mass calibration of redMaPPer galaxy clusters in dark energy survey science verification data, *Mon. Not. R. Astron. Soc.* **469**, 4899 (2017).
  - [13] E. Medezinski, M. Oguri, A. J. Nishizawa, J. S. Speagle, H. Miyatake, K. Umetsu, A. Leauthaud, R. Murata, R. Mandelbaum, C. Sifón, M. A. Strauss, S. Huang, M. Simet, N. Okabe, M. Tanaka, and Y. Komiyama, Source selection for cluster weak lensing measurements in the hyper Suprime-Cam survey, *Publ. Astron. Soc. Jpn.* **70**, 30 (2018).
  - [14] H. Miyatake, N. Battaglia, M. Hilton, E. Medezinski, A. J. Nishizawa, S. More, S. Aiola, N. Bahcall, J. R. Bond, E. Calabrese *et al.*, Weak-lensing mass calibration of ACTPol Sunyaev-Zel’dovich clusters with the hyper Suprime-Cam survey, *Astrophys. J.* **875**, 63 (2019).
  - [15] T. McClintock, T. N. Varga, D. Gruen, E. Rozo, E. S. Rykoff, T. Shin, P. Melchior, J. DeRose, S. Seitz, J. P. Dietrich, E. Sheldon, Y. Zhang, A. von der Linden, T. Jeltema, A. B. Mantz, A. K. Romer *et al.*, Dark energy survey year 1 results: Weak lensing mass calibration of redMaPPer galaxy clusters, *Mon. Not. R. Astron. Soc.* **482**, 1352 (2019).
  - [16] E. S. Rykoff, E. Rozo, M. T. Busha, C. E. Cunha, A. Finoguenov, A. Evrard, J. Hao, B. P. Koester, A. Leauthaud, B. Nord, M. Pierre, R. Reddick, T. Sadibekova, E. S. Sheldon, and R. H. Wechsler, redMaPPer. I. Algorithm and SDSS DR8 catalog, *Astrophys. J.* **785**, 104 (2014).
  - [17] J. P. Dietrich, Y. Zhang, J. Song, C. P. Davis, T. A. McKay, L. Baruah, M. Becker, C. Benoist, M. Busha, L. A. N. da Costa, J. Hao, M. A. G. Maia, C. J. Miller, R. Ogando, A. K. Romer, E. Rozo, E. Rykoff, and R. Wechsler, Orientation bias of optically selected galaxy clusters and its impact on stacked weak-lensing analyses, *Mon. Not. R. Astron. Soc.* **443**, 1713 (2014).
  - [18] M. Simet, T. McClintock, R. Mandelbaum, E. Rozo, E. Rykoff, E. Sheldon, and R. H. Wechsler, Weak lensing measurement of the Mass–Richness relation of SDSS redMaPPer clusters, *Mon. Not. R. Astron. Soc.* **466**, 3103 (2017).
  - [19] M. Costanzi, E. Rozo, M. Simet, Y. Zhang, A. E. Evrard, A. Mantz, E. S. Rykoff, T. Jeltema, D. Gruen, S. Allen *et al.*, Methods for cluster cosmology and application to the SDSS in preparation for DES Year 1 release, *Mon. Not. R. Astron. Soc.* **488**, 4779 (2019).
  - [20] T. M. C. Abbott, F. B. Abdalla, A. Alarcon, J. Aleksić, S. Allam, S. Allen, A. Amara, J. Annis, J. Asorey, S. Avila *et al.* (DES Collaboration), Dark energy survey year 1

- results: Cosmological constraints from galaxy clustering and weak lensing, *Phys. Rev. D* **98**, 043526 (2018).
- [21] A. Drlica-Wagner, I. Sevilla-Noarbe, E. S. Rykoff, R. A. Gruendl, B. Yanny, D. L. Tucker, B. Hoyle, A. C. Rosell, G. M. Bernstein, K. Bechtol *et al.*, Dark energy survey year 1 results: The photometric data set for cosmology, *Astrophys. J. Suppl. Ser.* **235**, 33 (2018).
- [22] J. E. Carlstrom, P. A. R. Ade, K. A. Aird, B. A. Benson, L. E. Bleem, S. Buseti, C. L. Chang, E. Chauvin, H.-M. Cho, T. M. Crawford *et al.*, The 10 meter south pole telescope, *Publ. Astron. Soc. Pac.* **123**, 568 (2011).
- [23] J. Annis, M. Soares-Santos, M. A. Strauss, A. C. Becker, S. Dodelson, X. Fan, J. E. Gunn, J. Hao, Ž. Ivezić, S. Jester *et al.*, The Sloan digital sky survey Coadd: 275 deg<sup>2</sup> of deep Sloan digital sky survey imaging on stripe 82, *Astrophys. J.* **794**, 120 (2014).
- [24] E. Rozo and E. S. Rykoff, redMaPPer II: X-ray and SZ performance benchmarks for the SDSS catalog, *Astrophys. J.* **783**, 80 (2014).
- [25] T. Sadibekova, M. Pierre, N. Clerc, L. Faccioli, R. Gastaud, J.-P. Le Fevre, E. Rozo, and E. Rykoff, The X-CLASS-redMaPPer galaxy cluster comparison. I. Identification procedures, *Astron. Astrophys.* **571**, A87 (2014).
- [26] E. Rozo, E. S. Rykoff, J. G. Bartlett, and J.-B. Melin, redMaPPer-III. A detailed comparison of the Planck 2013 and SDSS DR8 redMaPPer cluster catalogues, *Mon. Not. R. Astron. Soc.* **450**, 592 (2015).
- [27] A. Saro, S. Bocquet, E. Rozo, B. A. Benson, J. Mohr, E. S. Rykoff, M. Soares-Santos, L. Bleem, S. Dodelson, P. Melchior *et al.*, Constraints on the richness-mass relation and the optical-SZE positional offset distribution for SZE-selected clusters, *Mon. Not. R. Astron. Soc.* **454**, 2305 (2015).
- [28] E. Rozo, E. S. Rykoff, M. Becker, R. M. Reddick, and R. H. Wechsler, redMaPPer-IV. Photometric membership identification of red cluster galaxies with 1 per cent precision, *Mon. Not. R. Astron. Soc.* **453**, 38 (2015).
- [29] E. S. Rykoff, E. Rozo, D. Hollowood, A. Bermeo-Hernandez, T. Jeltema, J. Mayers, A. K. Romer, P. Rooney, A. Saro, C. V. Cervantes, R. H. Wechsler, H. Wilcox *et al.*, The RedMaPPer Galaxy cluster catalog from DES science verification data, *Astrophys. J. Suppl. Ser.* **224**, 1 (2016).
- [30] Y. Zhang, T. Jeltema, D. L. Hollowood, S. Everett, E. Rozo, A. Farahi, A. Bermeo, S. Bhargava, P. Giles, A. K. Romer *et al.*, Dark energy surveyed year 1 results: Calibration of cluster mis-centring in the redMaPPer catalogues, *Mon. Not. R. Astron. Soc.* **487**, 2578 (2019).
- [31] J. D. Cohn, A. E. Evrard, M. White, D. Croton, and E. Ellingson, Red-sequence cluster finding in the Millennium simulation, *Mon. Not. R. Astron. Soc.* **382**, 1738 (2007).
- [32] A. Farahi, A. E. Evrard, E. Rozo, E. S. Rykoff, and R. H. Wechsler, Galaxy cluster mass estimation from stacked spectroscopic analysis, *Mon. Not. R. Astron. Soc.* **460**, 3900 (2016).
- [33] Y. Zu, R. Mandelbaum, M. Simet, E. Rozo, and E. S. Rykoff, On the level of cluster assembly bias in SDSS, *Mon. Not. R. Astron. Soc.* **470**, 551 (2017).
- [34] P. Busch and S. D. M. White, Assembly bias and splashback in galaxy clusters, *Mon. Not. R. Astron. Soc.* **470**, 4767 (2017).
- [35] B. M. S. Erickson, C. E. Cunha, and A. E. Evrard, Influence of projection in cluster cosmology studies, *Phys. Rev. D* **84**, 103506 (2011).
- [36] M. Costanzi, E. Rozo, E. S. Rykoff, A. Farahi, T. Jeltema, A. E. Evrard, A. Mantz, D. Gruen, R. Mandelbaum, J. DeRose, T. McClintock, T. N. Varga, Y. Zhang, J. Weller, R. H. Wechsler, and M. Aguena, Modelling projection effects in optically selected cluster catalogues, *Mon. Not. R. Astron. Soc.* **482**, 490 (2019).
- [37] J. Zuntz, E. Sheldon, S. Samuroff, M. A. Troxel, M. Jarvis, N. MacCrann, D. Gruen, J. Prat, C. Sánchez, A. Choi *et al.*, Dark energy survey year 1 results: Weak lensing shape catalogues, *Mon. Not. R. Astron. Soc.* **481**, 1149 (2018).
- [38] E. S. Sheldon and E. M. Huff, Practical Weak-lensing shear measurement with metacalibration, *Astrophys. J.* **841**, 24 (2017).
- [39] E. Huff and R. Mandelbaum, Metacalibration: Direct self-calibration of biases in shear measurement, *arXiv:1702.02600*.
- [40] E. Sheldon, *NGMIX: Gaussian Mixture Models for 2D Images*, Astrophysics Source Code Library (2015), <http://ascl.net/1508.008>.
- [41] J. Zuntz, T. Kacprzak, L. Voigt, M. Hirsch, B. Rowe, and S. Bridle, IM3SHAPE: A maximum likelihood galaxy shear measurement code for cosmic gravitational lensing, *Mon. Not. R. Astron. Soc.* **434**, 1604 (2013).
- [42] B. T. P. Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, and M. S. S. Gill, GALSIM: The modular galaxy image simulation toolkit, *Astron. Comput.* **10**, 121 (2015).
- [43] N. Benítez, Bayesian photometric redshift estimation, *Astrophys. J.* **536**, 571 (2000).
- [44] D. Coe, N. Benítez, S. F. Sánchez, M. Jee, R. Bouwens, and H. Ford, Galaxies in the Hubble ultra deep field. I. Detection, multiband photometry, photometric redshifts, and morphology, *Astron. J.* **132**, 926 (2006).
- [45] M. Gatti, P. Vielzeuf, C. Davis, R. Cawthon, M. M. Rau, J. DeRose, J. De Vicente, A. Alarcon, E. Rozo, E. Gaztanaga, B. Hoyle, R. Miquel *et al.*, Dark energy survey year 1 results: Cross-correlation redshifts—methods and systematics characterization, *Mon. Not. R. Astron. Soc.* **477**, 1664 (2018).
- [46] C. Davis, M. Gatti, P. Vielzeuf, R. Cawthon, E. Rozo *et al.*, Dark energy survey year 1 results: Cross-correlation redshifts in the DES—Calibration of the weak lensing source redshift distributions, *arXiv:1710.02517*.
- [47] J. Prat, C. Sánchez, Y. Fang, D. Gruen, J. Elvin-Poole, N. Kokron, L. F. Secco, B. Jain, R. Miquel, N. MacCrann, M. A. Troxel *et al.*, Dark energy survey year 1 results: Galaxy-Galaxy lensing, *Phys. Rev. D* **98**, 042005 (2018).
- [48] B. Hoyle, D. Gruen, G. M. Bernstein, M. M. Rau, J. De Vicente, W. G. Hartley, E. Gaztanaga, J. DeRose, M. A. Troxel, C. Davis, A. Alarcon, N. MacCrann, J. Prat, C. Sánchez, E. Sheldon, R. H. Wechsler *et al.*, Dark energy survey year 1 results: Redshift distributions of the weak lensing source galaxies, *Mon. Not. R. Astron. Soc.* **478**, 592 (2018).
- [49] W. Hu and A. V. Kravtsov, Sample variance considerations for cluster surveys, *Astrophys. J.* **584**, 702 (2003).
- [50] W. Hu and J. D. Cohn, Likelihood methods for cluster dark energy surveys, *Phys. Rev. D* **73**, 067301 (2006).



- [51] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlöber, The large-scale bias of dark matter halos: Numerical calibration and model tests, *Astrophys. J.* **724**, 878 (2010).
- [52] C. Laigle, H. J. McCracken, O. Ilbert, B. C. Hsieh, I. Davidzon, P. Capak, G. Hasinger, J. D. Silverman, C. Pichon, J. Coupon *et al.*, The COSMOS2015 catalog: Exploring the  $1 < z < 6$  Universe with half a million galaxies, *Astrophys. J. Suppl. Ser.* **224**, 24 (2016).
- [53] T. N. Varga, J. DeRose, D. Gruen, T. McClintock, S. Seitz, E. Rozo, M. Costanzi, B. Hoyle, N. MacCrann, A. A. Plazas, E. S. Rykoff, M. Simet, A. von der Linden, R. H. Wechsler *et al.*, Dark energy survey year 1 results: Validation of weak lensing cluster member contamination estimates from  $P(z)$  decomposition, *Mon. Not. R. Astron. Soc.* **489**, 2511 (2019).
- [54] D. Gruen, S. Seitz, F. Brimiouille, R. Kosyra, J. Koppenhoefer, C. H. Lee, R. Bender, A. Riffeser, T. Eichner, T. Weidinger, and M. Bierschenk, Weak lensing analysis of SZ-selected clusters of galaxies from the SPT and Planck surveys, *Mon. Not. R. Astron. Soc.* **442**, 1507 (2014).
- [55] D. Gruen, S. Seitz, M. R. Becker, O. Friedrich, and A. Mana, Cosmic variance of the galaxy cluster weak lensing signal, *Mon. Not. R. Astron. Soc.* **449**, 4264 (2015).
- [56] M. A. Henson, D. J. Barnes, S. T. Kay, I. G. McCarthy, and J. Schaye, The impact of baryons on massive galaxy clusters: Halo structure and cluster mass estimates, *Mon. Not. R. Astron. Soc.* **465**, 3361 (2017).
- [57] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz, Toward a halo mass function for precision cosmology: The limits of universality, *Astrophys. J.* **688**, 709 (2008).
- [58] J. DeRose, R. H. Wechsler, J. L. Tinker, M. R. Becker, Y.-Y. Mao, T. McClintock, S. McLaughlin, E. Rozo, and Z. Zhai, The Aemulus project I: Numerical simulations for precision cosmology, *Astrophys. J.* **875**, 69 (2019).
- [59] M. Costanzi, F. Villaescusa-Navarro, M. Viel, J.-Q. Xia, S. Borgani, E. Castorina, and E. Sefusatti, Cosmology with massive neutrinos III: The halo mass function and an application to galaxy clusters, *J. Cosmology Astropart. Phys.* **12** (2013) 012.
- [60] W. Cui, S. Borgani, and G. Murante, The effect of active galactic nuclei feedback on the halo mass function, *Mon. Not. R. Astron. Soc.* **441**, 1769 (2014).
- [61] M. Velliscig, M. P. van Daalen, J. Schaye, I. G. McCarthy, M. Cacciato, A. M. C. Le Brun, and C. Dalla Vecchia, The impact of galaxy formation on the total mass, mass profile and abundance of haloes, *Mon. Not. R. Astron. Soc.* **442**, 2641 (2014).
- [62] S. Bocquet, A. Saro, K. Dolag, and J. J. Mohr, Halo mass function: Baryon impact, fitting formulae, and implications for cluster cosmology, *Mon. Not. R. Astron. Soc.* **456**, 2361 (2016).
- [63] V. Springel, R. Pakmor, A. Pillepich, R. Weinberger, D. Nelson, L. Hernquist, M. Vogelsberger, S. Genel, P. Torrey, F. Marinacci, and J. Naiman, First results from the IllustrisTNG simulations: Matter and galaxy clustering, *Mon. Not. R. Astron. Soc.* **475**, 676 (2018).
- [64] J. DeRose, R. H. Wechsler, M. R. Becker, M. T. Busha, E. S. Rykoff, N. MacCrann, B. Erickson, A. E. Evrard, A. Kravtsov, D. Gruen *et al.*, The buzzard flock: Dark energy survey synthetic sky catalogs, [arXiv:1901.02401](https://arxiv.org/abs/1901.02401).
- [65] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, The 6dF Galaxy survey: Baryon acoustic oscillations and the local Hubble constant, *Mon. Not. R. Astron. Soc.* **416**, 3017 (2011).
- [66] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera, The clustering of the SDSS DR7 main Galaxy sample—I. A 4 per cent distance measure at  $z = 0.15$ , *Mon. Not. R. Astron. Soc.* **449**, 835 (2015).
- [67] S. Alam *et al.*, The clustering of galaxies in the completed SDSS-III Baryon oscillation spectroscopic survey: Cosmological analysis of the DR12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [68] D. M. Scolnic *et al.*, The complete light-curve sample of spectroscopically confirmed SNe Ia from Pan-STARRS1 and cosmological constraints from the combined Pantheon sample, *Astrophys. J.* **859**, 101 (2018).
- [69] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC hammer, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [70] J. Goodman and J. Weare, Ensemble samplers with affine invariance, *Commun. Appl. Math. Comput. Sci.* **5**, 65 (2010).
- [71] M. Hasselfield, M. Hilton, T. A. Marriage, G. E. Addison, L. F. Barrientos, N. Battaglia, E. S. Battistelli, J. R. Bond, D. Crichton, S. Das *et al.*, The Atacama cosmology telescope: Sunyaev-Zel'dovich selected galaxy clusters at 148 GHz from three seasons of data, *J. Cosmology Astropart. Phys.* **07** (2013) 008.
- [72] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts, *Astron. Astrophys.* **594**, A24 (2016).
- [73] E. van Uitert *et al.*, KiDS+GAMA: Cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing, and angular clustering, *Mon. Not. R. Astron. Soc.* **476**, 4662 (2018).
- [74] S. Joudaki *et al.*, KiDS-450 + 2dFLenS: Cosmological parameter constraints from weak gravitational lensing tomography and overlapping redshift-space galaxy clustering, *Mon. Not. R. Astron. Soc.* **474**, 4894 (2018).
- [75] H. Hildebrandt *et al.*, KiDS + VIKING-450: Cosmic shear tomography with optical + infrared data, *Astron. Astrophys.* **633**, A69 (2020).
- [76] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [77] T. Charnock, R. A. Battye, and A. Moss, Planck data versus large scale structure: Methods to quantify discordance, *Phys. Rev. D* **95**, 123535 (2017).
- [78] E. Rozo, R. H. Wechsler, E. S. Rykoff, J. T. Annis, M. R. Becker, A. E. Evrard, J. A. Frieman, S. M. Hansen, J. Hao, D. E. Johnston, B. P. Koester, T. A. McKay, E. S. Sheldon, and D. H. Weinberg, Cosmological constraints from the Sloan digital sky survey maxBCG cluster catalog, *Astrophys. J.* **708**, 645 (2010).

- [79] R. Murata, M. Oguri, T. Nishimichi, M. Takada, R. Mandelbaum, S. More, M. Shirasaki, A. J. Nishizawa, and K. Osato, The mass-richness relation of optically selected clusters from weak gravitational lensing and abundance with Subaru HSC first-year data, *Publ. Astron. Soc. Jpn.* **71**, 107 (2019).
- [80] L. E. Bleem *et al.*, Galaxy clusters discovered via the Sunyaev-Zel'dovich effect in the 2500-square-degree SPT-SZ survey, *Astrophys. J. Suppl. Ser.* **216**, 27 (2015).
- [81] L. E. Bleem, S. Bocquet, B. Stalder, M. D. Gladders *et al.*, The SPTpol extended cluster survey, *Astrophys. J. Suppl. Ser.* **247**, 25 (2020).
- [82] D. L. Hollowood, T. Jeltema, X. Chen, A. Farahi, A. Evrard, S. Everett, E. Rozo, E. Rykoff, R. Bernstein, A. Bermeo, L. Eiger, P. Giles, H. Israel, R. Michel, R. Noorali, K. Romer, P. Rooney, and M. Spletstoesser, Chandra follow-up of the SDSS DR8 redMaPPer catalog using the MATCha pipeline, *Astrophys. J. Suppl. Ser.* **244**, 22 (2019).
- [83] E. J. Baxter, S. Raghunathan, T. M. Crawford, P. Fosalba, Z. Hou, G. P. Holder, Y. Omori, S. Patil, E. Rozo *et al.*, A measurement of CMB cluster lensing with SPT and DES year 1 data, *Mon. Not. R. Astron. Soc.* **476**, 2674 (2018).
- [84] S. Raghunathan, S. Patil, E. Baxter, B. A. Benson, L. E. Bleem, T. M. Crawford, G. P. Holder, T. McClintock, C. L. Reichardt, T. N. Varga, N. Whitehorn *et al.*, A Detection of CMB-Cluster Lensing Using Polarization Data from SPTpol, *Phys. Rev. Lett.* **123**, 181301 (2019).
- [85] R. Garcia and E. Rozo, Halo exclusion criteria impacts halo statistics, *Mon. Not. Roy. Astron. Soc.* **489**, 4170 (2019).
- [86] A. Leauthaud *et al.*, Lensing is low: Cosmology, galaxy formation or new physics?, *Mon. Not. R. Astron. Soc.* **467**, 3024 (2017).
- [87] J. E. Geach and J. A. Peacock, Cluster richness-mass calibration with cosmic microwave background lensing, *Nat. Astron.* **1**, 795 (2017).
- [88] R. Capasso *et al.*, Mass calibration of the CODEX cluster sample using SPIDERS spectroscopy I. The richness-mass relation, *Mon. Not. R. Astron. Soc.* **486**, 1594 (2019).
- [89] N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, R. B. Barreiro, M. Bartelmann, J. G. Bartlett *et al.* (Planck Collaboration), Planck early results. XII. Cluster Sunyaev-Zeldovich optical scaling relations, *Astron. Astrophys.* **536**, A12 (2011).
- [90] B. P. Koester, T. A. McKay, J. Annis, R. H. Wechsler, A. E. Evrard, E. Rozo, L. Bleem, E. S. Sheldon, and D. Johnston, MaxBCG: A red-sequence Galaxy cluster finder, *Astrophys. J.* **660**, 221 (2007).
- [91] E. Rozo, J. G. Bartlett, A. E. Evrard, and E. S. Rykoff, Closing the loop: a self-consistent model of optical, x-ray and Sunyaev-Zel'dovich scaling relations for clusters of Galaxies, *Mon. Not. R. Astron. Soc.* **438**, 78 (2014).
- [92] A. Palmese, J. Annis, J. Burgad, A. Farahi, M. Soares-Santos, B. Welch, M. da Silva Pereira, H. Lin, S. Bhargava, D. L. Hollowood *et al.*, Stellar mass as a galaxy cluster mass proxy: Application to the dark energy survey redMaPPer clusters, *Mon. Not. R. Astron. Soc.* **493**, 4591 (2020).
- [93] C. Bradshaw, A. Leauthaud, A. Hearin, S. Huang, and P. Behroozi, Physical correlations of the scatter between Galaxy mass, stellar content, and halo mass, *Mon. Not. R. Astron. Soc.* **493**, 337 (2020).
- [94] A. E. Evrard, P. Arnault, D. Huterer, and A. Farahi, A model for multiproperty galaxy cluster statistics, *Mon. Not. R. Astron. Soc.* **441**, 3562 (2014).
- [95] Y. Noh and J. D. Cohn, Disentangling correlated scatter in cluster mass measurements, *Mon. Not. R. Astron. Soc.* **426**, 1829 (2012).
- [96] A. Gelman and D. B. Rubin, Inference from iterative simulation using multiple sequences, *Stat. Sci.* **7**, 457 (1992).
- [97] T. Sunayama, Y. Park, M. Takada, Y. Kobayashi, T. Nishimichi, T. Kurita, S. More, M. Oguri, and K. Osato, The impact of projection effects on cluster observables: stacked lensing and projected clustering, [arXiv:2002.03867](https://arxiv.org/abs/2002.03867).