Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **502,** 6010–6031 (2021) doi:10.1093/mnras/stab357
Advance Access publication 2021 February 15

Dark energy survey year 1 results: Constraining baryonic physics in the Universe

Hung-Jin Huang [1,2★] Tim Eifler [1] Rachel Mandelbaum [2] Gary M. Bernstein [3] Anqi Chen,[4] Ami Choi,[5] Juan García-Bellido [6] Dragan Huterer,[4] Elisabeth Krause,[1] Eduardo Rozo,[7] Sukhdeep Singh,[8] Sarah Bridle,[9] Joseph DeRose,[8,10] Jack Elvin-Poole,[5,11] Xiao Fang [1] Oliver Friedrich,[12] Marco Gatti,[13] Enrique Gaztanaga [14,15] Daniel Gruen [16,17,18] Will Hartley,[19,20,21] Ben Hoyle [22,23] Mike Jarvis,[3] Niall MacCrann [5,11] Vivian Miranda,[1] Markus Rau [2] Judit Prat [24] Carles Sánchez [3] Simon Samuroff [2] Michael Troxel [25] Joe Zuntz [26] Tim Abbott,[27] Michel Aguena,[28,29] James Annis,[30] Santiago Avila [6] Matthew Becker [31] Emmanuel Bertin [32,33] David Brooks,[20] David Burke,[17,18] Aurelio Carnero Rosell [34,35] Matias Carrasco Kind [36,37] Jorge Carretero,[13] Francisco Javier Castander,[14,15] Luiz da Costa,[29,38] Juan De Vicente [39] Jörg Dietrich,[40] Peter Doel,[20] Spencer Everett,[10] Brenna Flaugher,[30] Pablo Fosalba [14,15] Josh Frieman,[30,41] Robert Gruendl,[36,37] Gaston Gutierrez,[30] Samuel Hinton [42] Klaus Honscheid,[5,11] David James,[43] Kyler Kuehn,[44,45] Ofer Lahav,[20] Marcos Lima,[28,29] Marcio Maia,[29,38] Jennifer Marshall,[46] Felipe Menanteau,[36,37] Ramon Miquel,[13,47] Francisco Paz-Chinchón,[12,37] Andrés Plazas Malagón [48] Kathy Romer,[49] Aaron Roodman,[17,18] Eusebio Sanchez,[39] Vic Scarpine,[30] Santiago Serrano,[14,15] Ignacio Sevilla,[39] Mathew Smith [50] Marcelle Soares-Santos,[51] Eric Suchyta [52] Molly Swanson,[37] Gregory Tarle,[4] Diehl H. Thomas,[30] and Jochen Weller[22,23] The DES Collaboration

Affiliations are listed at the end of the paper

Accepted 2021 February 4. Received 2021 January 11; in original form 2020 July 30

ABSTRACT

Measurements of large-scale structure are interpreted using theoretical predictions for the matter distribution, including potential impacts of baryonic physics. We constrain the feedback strength of baryons jointly with cosmology using weak lensing and galaxy clustering observables (3×2pt) of Dark Energy Survey (DES) Year 1 data in combination with external information from baryon acoustic oscillations (BAO) and Planck cosmic microwave background polarization. Our baryon modelling is informed by a set of hydrodynamical simulations that span a variety of baryon scenarios; we span this space via a Principal Component (PC) analysis of the summary statistics extracted from these simulations. We show that at the level of DES Y1 constraining power, one PC is sufficient to describe the variation of baryonic effects in the observables, and the first PC amplitude (Q_1) generally reflects the strength of baryon feedback. With the upper limit of Q_1 prior being bound by the Illustris feedback scenarios, we reach ~ 20 per cent improvement in the constraint of $S_8 = \sigma_8(\Omega_m/0.3)^{0.5} = 0.788^{+0.018}_{-0.021}$ compared to the original DES 3×2pt analysis. This gain is driven by the inclusion of small-scale cosmic shear information down to 2.5 arcmin, which was excluded in previous DES analyses that did not model baryonic physics. We obtain $S_8 = 0.781^{+0.014}_{-0.015}$ for the combined DES Y1+Planck EE+BAO analysis with a non-informative Q_1 prior. In terms of the baryon constraints, we measure $Q_1 = 1.14^{+2.20}_{-2.80}$ for DES Y1 only and $Q_1 = 1.42^{+1.63}_{-1.48}$ for DESY1+Planck EE+BAO, allowing us to exclude one of the most extreme AGN feedback hydrodynamical scenario at more than 2σ.

Key words: cosmological parameters – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Understanding the composition and evolution of our Universe has been a central science endeavour in the astronomical community. Ongoing wide-field imaging surveys such as the Dark Energy Survey (DES;[1] Krause et al. 2017; Abbott et al. 2018, 2019; Troxel et al. 2018a), the Kilo-Degree Survey (KiDS;[2] van Uitert et al. 2018;

★ E-mail: hungjinh@email.arizona.edu

[1] www.darkenergysurvey.org/
[2] http://www.astro-wise.org/projects/KIDS/

Kuijken et al. 2019; Hildebrandt et al. 2020), and the Hyper Suprime Cam Subaru Strategic Program (HSC;[3] Mandelbaum et al. 2018; Hikage et al. 2019; Hamana et al. 2020) have collected a wealth of cosmological data over the past years that can be used to explore fundamental questions such as the underlying physics of cosmic acceleration, the mass and number of neutrino species, and the interplay of dark and luminous matter.

The cosmological information is bound to increase significantly in the near future with analysis of the full DES, KiDS, and HSC data sets, and even more so in the early 2020s with the advent of Stage IV surveys such as the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST;[4] Ivezić et al. 2019), *Euclid*[5] (Laureijs et al. 2011), the Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (SPHEREx;[6] Doré et al. 2014), and the *Nancy Grace Roman Space Telescope* (*WFIRST*;[7] Spergel et al. 2015; Eifler et al. 2020a,b).

The increased cosmological information encoded in these data sets will require a new level in accuracy of modelling cosmological observables. One of the fundamental quantities for making theoretical predictions is the matter power spectrum $P_\delta(k, z)$, which quantifies the amount of matter clustering at the second-order level and its evolution as a function of time. Previous studies have estimated that $P_\delta(k, z)$ needs to be predicted to ~ 1 per cent level out to $k \lesssim 10\,h\mathrm{Mpc}^{-1}$ for the future era of Stage IV cosmological experiments (e.g. Huterer & Takada 2005; Eifler 2011; Hearin, Zentner & Ma 2012). To quantify the non-linear evolution of the density field at the required precision, significant computational resources have been devoted to building power spectrum emulators with N-body dark-matter-only (DMO) simulations (e.g. Heitmann et al. 2010, 2014; DeRose et al. 2019). However, baryonic effects such as feedback and cooling mechanisms redistribute matter, causing uncertainties in $P_\delta(k, z)$ at the level of tens of per cent (e.g. van Daalen et al. 2011; Chisari et al. 2018; van Daalen, McCarthy & Schaye 2020) for $k \gtrsim 5\,h\mathrm{Mpc}^{-1}$.

Adopting mitigation schemes to account for uncertainties of baryons is crucial to assure the robustness of cosmological analyses. The most straightforward way is to exclude data points for which the fractional contributions from potential systematic uncertainties are non-negligible given the covered model flexibility. For the DES Y1 cosmic shear analysis, conservative scale cuts are applied to ensure the level of baryon contamination to be within 2 per cent (Troxel et al. 2018a).

Methods have been proposed to reduce the sensitivity to small scales in the data. By cutting the most extreme peaks in the density fields, the derived summary statistics become less sensitive to the non-linear regime, as proposed in the peak clipping technique (Simpson et al. 2011; Simpson, Heavens & Heymans 2013; Giblin et al. 2018). By designing special weighting functions to filter out the contributions of small-scale modes in observables of cosmic shear, the k-cut (Taylor, Bernardeau & Kitching 2018) and x-cut (Taylor, Bernardeau & Huff 2020) cosmic shear methods provide new summary statistics with reduced sensitivity to baryonic effects on the matter power spectrum. Also, the COSEBIs (Schneider, Eifler & Krause 2010) method is designed to separate E/B modes from $\xi_\pm$ in a finite angular interval, which makes its summary statistics less sensitive to small-scale physical effects compared with $\xi_\pm$ given a fixed angular range (Asgari et al. 2020).

However, including small-scale information with models for baryonic effects not only provides the potential to increase the statistical power of constraints on cosmology, but also offers a mechanism to quantify the effects of baryons on the matter power spectrum using real data. A number of methods have been proposed to model baryonic effects (see Chisari et al. 2019 for a review). One class of methods is to employ the halo model (Peacock & Smith 2000; Seljak 2000), based on the principle that the main contribution of baryons is to modify halo density profiles in the one-halo regime (see e.g. Rudd, Zentner & Kravtsov 2008; Velliscig et al. 2014; Mummery et al. 2017). Within the NFW (Navarro–Frenk–White; Navarro, Frenk & White 1996) profile, a straightforward option is to vary parameters related to the halo concentration to perform baryon marginalization (Zentner, Rudd & Hu 2008; Zentner et al. 2013). Besides the degree of freedom provided via halo concentration, extra parameters are added offering the complexity to account for the effect of halo bloating induced by baryonic feedback in `HMCODE` (Mead et al. 2015, 2016), and for the inner halo core formation induced by the cooling effect of baryons (Copeland, Taylor & Hall 2018). `HMCODE` has been applied in several weak lensing analyses to mitigate baryonic effects, for example in data sets of CFHTLenS (Joudaki et al. 2017), DES Science Verification (MacCrann et al. 2017), KiDS-450 (Hildebrandt et al. 2017; Yoon & Jee 2020), and DLS (Yoon et al. 2019). Even more sophisticated halo model frameworks provide descriptions of the radial distributions of the stellar, gas, and dark matter components within haloes, and parametrize the baryonic effects in more physically motivated quantities (Semboloni et al. 2011; Semboloni, Hoekstra & Schaye 2013; Mohammed et al. 2014; Schneider & Teyssier 2015; Debackere, Schaye & Hoekstra 2020).

Another category of approaches to modelling baryonic effects is through empirical modelling, where the functional form of the fitting formula is calibrated based on hydrodynamical simulations. Parametric forms are designed with the flexibility to model the behaviour of the power spectrum ratio between paired hydrodynamical and DMO simulations ($P_{\delta,\,\mathrm{hydro}}(k)/P_{\delta,\,\mathrm{DMO}}(k)$) for the Horizon-AGN hydrosimulation in Chisari et al. (2018), and for the nine scenarios in the OWLS simulation suites as detailed in Harnois-Déraps et al. (2015), which is adopted in HSC Y1 cosmic shear analysis to account for baryonic effects (Hikage et al. 2019). Recently van Daalen et al. (2020) derive a formulation which provides even wider applications for hydrodynamical scenarios accumulated over the past ten years. Alternatively, Eifler et al. (2015) proposed performing principal component analyses (PCA) using the cosmic shear model vectors extracted from the hydrodynamical simulations, and use a few dominant principal components (PCs) as a flexible basis to span the range of baryon uncertainties for the survey-specific summary statistic.

Going beyond modelling summary statistics, there are approaches focusing on post-processing the output DMO simulations. The 'baryonification' model contains prescriptions for the density profiles of the stellar, gas, and the redistributed DM components to correct the particle positions in DMO simulations so as to more accurately approximate what they would look like in the presence of baryons (Schneider & Teyssier 2015; Schneider et al. 2019; Aricò et al. 2020). Dai, Feng & Seljak (2018) propose using the potential gradient descent model to displace particles to improve the modelling of non-linear matter distribution.

In this paper, we aim to utilize the information from small-scale cosmic shear data to place constraints on the strength of baryonic effects, and to compare the results with existing hydrodynamical

[3] http://www.naoj.org/Projects/HSC/HSCProject.html
[4] https://www.lsst.org/
[5] https://sci.esa.int/web/euclid
[6] http://spherex.caltech.edu/
[7] https://wfirst.gsfc.nasa.gov/

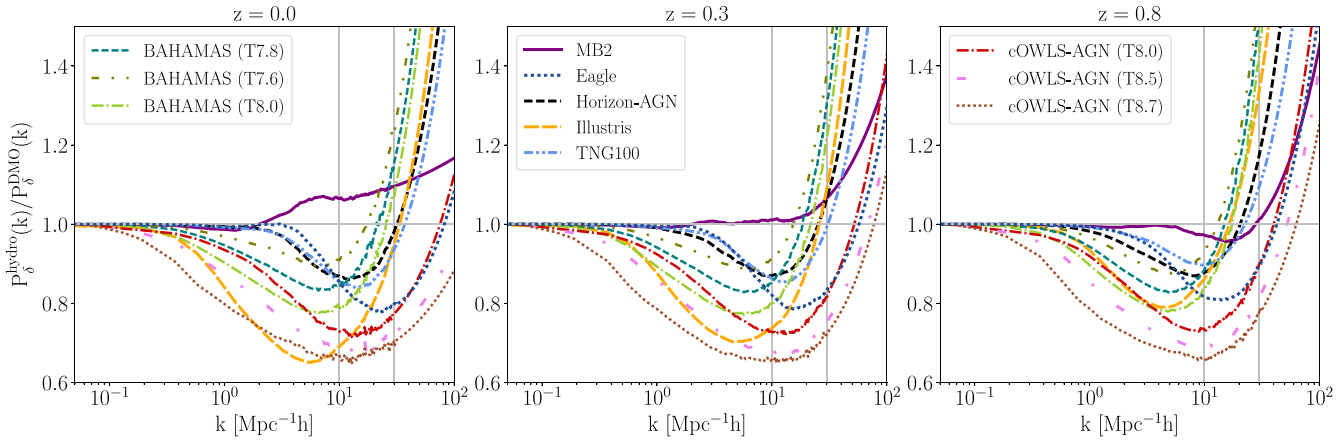

Figure 1. Baryonic effects on the 3D matter power spectrum at different redshifts. We plot the power spectrum ratio for 11 hydrodynamical simulations with respect to their corresponding DMO simulation setting at the same initial condition: IllustrisTNG, MB2, Eagle, Horizon-AGN, Illustris, the cosmo-OWLS sets, and the BAHAMAS sets, at redshifts 0.0, 0.3, 0.8. The grey vertical lines delineate regions where the data points come from direct measurement ($k < 30\,h$ Mpc^{-1}) and from extrapolation ($k > 30\,h$ Mpc^{-1}) with a quadratic spline fit based on data points at $k \in [10, 30]\,h$ Mpc^{-1}.

simulations. We will also explore the potential for achieving more precise cosmological constraints with the inclusion of small-scale data. We adopt the PCA baryon mitigation framework (Eifler et al. 2015) to perform our analyses. In Huang et al. (2019), we have validated and improved the performance of the PCA method using simulated analyses of cosmic shear mock data under an LSST-like survey configuration. Here we delve into its application to the observational data of DES Y1, which includes two-point correlations of cosmic shear, galaxy–galaxy lensing and galaxy clustering. With the ability to model small-scale cosmic shear, we push the cosmic shear observables down to 2.5 arcmin and perform a combined analysis with galaxy–galaxy lensing and galaxy clustering data (subjected to the original conservative Y1 scale cuts).

We begin with an overview of the data products, the theoretical modelling, and the analysis approaches in Section 2. We describe the design and validation of our pipeline in Section 3. We employ simulated likelihood analyses to understand and validate our pipeline performance, before we unblind and perform analyses of the real DES Y1 data. We present our main cosmology results in Section 4, followed by our constraints on baryonic effects in Section 5. Finally, we conclude in Section 6.

2 DATA, THEORY, AND ANALYSIS

2.1 Data

2.1.1 Observational data

In this work, we use the DES Y1 3 × 2pt data vector[8] which is computed using the METACALIBRATION (Huff & Mandelbaum 2017; Sheldon & Huff 2017; Zuntz et al. 2018) shape catalogue as the source sample for cosmic shear (Troxel et al. 2018a), and the redMaGiC (Rozo et al. 2016) sample as the lens population for galaxy–galaxy lensing (Prat et al. 2018) and galaxy clustering (Elvin-Poole et al. 2018) measurements. The photometric redshift measurement and calibration are described in Hoyle et al. (2018), Gatti et al. (2018), Davis et al. (2017).

[8]The publicly released 3 × 2pt data vector and its associated covariance matrix, `2pt_NG_mcal_1110.fits`, can be downloaded at https://des.ncsa.illinois.edu/releases/y1a1/key-products

The DES Y1 source galaxies are divided into four tomographic bins ranging from $z = 0.2$ to 1.3, resulting in 10 auto- and cross-correlations of cosmic shear for ξ_+ and ξ_-, respectively. The lens galaxies are placed in five tomographic bins ranging from $z = 0.15$ to 0.9, resulting in 20 tomographic cross-correlation bins between lens and source samples for galaxy–galaxy lensing, and five autocorrelations for galaxy clustering. Each of the correlation function statistics is measured using `treecorr` (Jarvis, Bernstein & Jain 2004) in 20 log-spaced bins of angular separation 2.5 arcmin $< \theta <$ 250 arcmin.

Conservative scale cuts are applied to the raw 3 × 2pt data vector in the original DES Y1 key cosmological analysis to avoid biases due to modelling uncertainties on small scales (Abbott et al. 2018).

For cosmic shear, scale cuts are determined by contaminating the $\xi^\pm$ model vector according to the OWLS-AGN scenario (Schaye et al. 2010), which has the same baryonic feature as the cosmo-OWLS AGN scenario shown in the red curves in Fig. 1, and removing data points that have a fractional contribution of baryons exceeding 2 per cent (Troxel et al. 2018a). For galaxy–galaxy lensing and galaxy clustering, the scale cuts are defined using a specific comoving scales of $(R_{\rm ggl}, R_{\rm clustering}) = (12, 8)\,{\rm Mpc}\,h^{-1}$ to avoid parameter biases due to non-linear biasing or non-locality of γ_t, and converted to their corresponding angular scales in each tomographic bin (Krause et al. 2017). After scale cuts are applied, there are a total of 457 elements for the fiducial DES Y1 3 × 2pt cosmological analysis (Abbott et al. 2018).

In this analysis, we utilize the DES Y1 cosmic shear correlation function measurements down to scales of 2.5 arcmin. Together with the galaxy–galaxy lensing and galaxy clustering measurements (subjected to the original DES Y1 scale cuts), our extended 3 × 2pt data vector has a total of 630 data points (400 elements for cosmic shear, 176 elements for galaxy–galaxy lensing, and 54 elements for galaxy clustering).

2.1.2 Hydrodynamical simulation data and power spectrum

In order to build baryon mitigation models with sufficient flexibility, we rely on a large variety of hydrodynamical simulations: MassiveBlack-II (MB2; Khandai et al. 2015; Tenneti et al. 2015), Horizon-AGN (Dubois et al. 2014), Eagle (Schaye et al. 2015), Illustris (Genel et al. 2014; Vogelsberger et al. 2014), IllustrisTNG

(Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018), three cosmo-OWLS simulations (cOWLS; Le Brun et al. 2014) with their minimum active galactic nucleus (AGN) heating temperatures $\Delta T_{\rm heat}$ being set at $10^{8.0}$, $10^{8.5}$, $10^{8.7}$, and three BAHAMAS scenarios (McCarthy et al. 2017) with their $\Delta T_{\rm heat} = 10^{7.6}$, $10^{7.8}$, $10^{8.0}$. Here $\Delta T_{\rm heat}$ is the most dominant subgrid physical parameter controlling the strength of AGN feedback in the cosmo-OWLS and the BAHAMAS simulation sets. Black holes are storing their feedback energy until it is large enough to heat the a certain number of surrounding particles by $\Delta T_{\rm heat}$.

Fig. 1 shows the effects of baryonic physics on the 3D matter power spectra for different hydrodynamical simulations, displayed as the ratio of these power spectra with respect to the power spectra for the corresponding dark matter only (DMO) simulations with the same initial conditions. On small scales, the effects of baryons show large variations, and have different redshift evolution histories across simulations. On large scales, we expect the power spectrum ratios to converge to unity because of diminishing baryonic effects, and because of the cosmic variance fluctuations being cancelled when taking ratios of power spectra for pairs of simulations with identical initial conditions. In Appendix B of Huang et al. 2019 (hereafter H19), we have discussed the convergence of power spectrum ratios in detail and provide an upper limit for their uncertainties due to cosmic variance.

We have used the power spectrum ratio for MB2, ILLUSTRIS, Eagle from H19. We extracted power spectrum measurements from the publicly released IllustrisTNG100 snapshot data (Nelson et al. 2019) and added the corresponding baryonic scenario to our power spectrum library. The Horizon-AGN $P_\delta(k)$ data are computed in Chisari et al. (2018). The cosmo-OWLS and BAHAMAS $P_\delta(k)$ sets are taken from the power spectra library released by van Daalen et al. (2020). Specifically, as listed in table 1 of van Daalen et al. (2020), for the cosmo-OWLS baryonic scenario sets, we use the $P_\delta(k)$ data from files:

(i) *AGN_Mseed800_WMAP7_L100N512*,
(ii) *AGN_Mseed800_Theat 8p5_WMAP7_L100N512*,
(iii) *AGN_Mseed800_Theat 8p7_WMAP7_L100N512*,

for the BAHAMAS sets, we use files:

(i) *AGN_CALIB_nu0_WMAP9_L400N1024*,
(ii) *AGN_CALIB_Theat_7p6_nu0_WMAP9_L400N1024*,
(iii) *AGN_CALIB_Theat_8p0_nu0_WMAP9_L400N1024*.

We make a slight adjustment to the power spectrum ratios. At larger scales, the raw $P_\delta(k)$ ratios for Horizon-AGN and cosmo-OWLS are observed to have subtle ($\lesssim 1$ per cent) excesses above unity towards large scales (e.g. see fig. 5 of Chisari et al. 2018). As discussed in Appendix B of van Daalen et al. (2020), this large-scale excess of power originates from details of the simulation setup between pairs of hydrodynamical and DMO simulations, for which their transfer functions and the number of particles often differ. Given that this sub-percent level offset is due to artefacts, we correct for this power mismatch by re-scaling the DMO power spectra using the linear growth factor, such that the ratio between $P_\delta^{\rm hydro}$ and $P_\delta^{\rm DMO}$ asymptotically approaches one on large scales.

On scales above $k > 30\,{\rm Mpc}^{-1}\,h$, we perform extrapolation by fitting a quadratic spline curve to data points at $k \in [10, 30]\,{\rm Mpc}^{-1}\,h$ to capture the power boosting from the effect of cooling. As discussed in Appendix B of H19, we argue that our extrapolation approach more accurately captures cooling effects compared to simply adopting the raw ratio as computed from the simulations. This is supported by comparing both methods to power spectrum ratios derived from higher resolution simulations.

Table 1. Parameters and priors used to run the likelihood analyses. Flat(a, b) denotes a flat prior in the range given while Gauss(μ, σ) is a Gaussian prior with mean μ and width σ. The third column summarizes the fiducial parameter values we used to generate mock data vectors and to construct PCs. The fiducial values are chosen to be consistent with the posterior constraints from the fiducial ΛCDM model of DES Y1 3 × 2pt analyses (Abbott et al. 2018). The fiducial photo-z and shear calibration parameters are set at the peak of the Gaussian prior for the purpose of running likelihood simulations.

Parameter	Prior	Fiducial value
	Cosmology	
$\Omega_{\rm m}$	Flat (0.1, 0.9)	0.3
$A_{\rm s}$	Flat (5×10^{-10}, 5×10^{-9})	2.19×10^{-9}
$n_{\rm s}$	Flat (0.87, 1.07)	0.97
$\Omega_{\rm b}$	Flat (0.03, 0.07)	0.048
$\Omega_\nu h^2$	baseline : Flat (5×10^{-4}, 0.0013)	0.00083
	Y1 fiducial : Flat (5×10^{-4}, 0.01)	–
h	Flat (0.55, 0.91)	0.69
	Lens galaxy bias	
$b_{\rm g}^1$	Flat (0.8, 3.0)	1.53
$b_{\rm g}^2$	Flat (0.8, 3.0)	1.71
$b_{\rm g}^3$	Flat (0.8, 3.0)	1.70
$b_{\rm g}^4$	Flat (0.8, 3.0)	2.05
$b_{\rm g}^5$	Flat (0.8, 3.0)	2.14
	Lens photo-z shift	
$\Delta z_{\rm l}^1$	Gauss (0.008, 0.007)	0.008
$\Delta z_{\rm l}^2$	Gauss (−0.005, 0.007)	−0.005
$\Delta z_{\rm l}^3$	Gauss (0.006, 0.006)	0.006
$\Delta z_{\rm l}^4$	Gauss (0.0, 0.01)	0.0
$\Delta z_{\rm l}^5$	Gauss (0.0, 0.01)	0.0
	Source photo-z shift	
$\Delta z_{\rm s}^1$	Gauss (−0.001, 0.016)	−0.001
$\Delta z_{\rm s}^2$	Gauss (−0.019, 0.013)	−0.019
$\Delta z_{\rm s}^3$	Gauss (+0.009, 0.011)	0.009
$\Delta z_{\rm s}^4$	Gauss (−0.018, 0.022)	−0.018
	Shear calibration (METACALIBRATION)	
m^1	Gauss (0.012, 0.023)	0.012
m^2	Gauss (0.012, 0.023)	0.012
m^3	Gauss (0.012, 0.023)	0.012
m^4	Gauss (0.012, 0.023)	0.012
	Intrinsic alignment	
$A_{\rm IA}$	Flat (−5, 5)	0.45
$\eta_{\rm IA}$	Flat (−5, 5)	−1.0
	Baryon PC amplitude	
Q_1	baseline: Flat (−3, 12) informative: Flat (0, 4)	
Q_2	Flat (−2.5, 2.5)	

2.1.3 *Mock data vectors*

In order to validate our baryon mitigation pipeline, we generate three mock data vectors to conduct simulated likelihood analyses: a pure theoretical data vector derived from our analysis pipeline (COSMOLIKE) with the fiducial parameters shown in Table 1 (we refer to this mock data vector as the DMO scenario hereafter), and

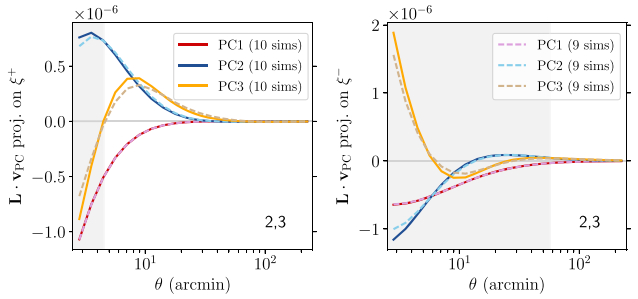

Figure 2. The principal components used in our baryon model (equation 19). Here we show the first three $\mathbf{L} \cdot \mathbf{v}_{\rm PC}$ components projected on the cosmic shear correlation functions in the cross tomographic bin (2,3). The solid curves indicate PCs constructed based on 10 hydrodynamical scenarios, which are used when analysing the DES data, and validating our pipeline on mock data constructed from the Illustris and the DMO scenarios. The dashed curves are constructed from 9 hydrodynamical scenarios, which are used when validating on the Eagle mock data (see Section 2.3.1 for detail). The grey shaded backgroud regions highlight the angular scales that excluded in the original Y1 cosmic shear analysis. In this work, we include these small-scale cosmic shear data points, and use the PCs as flexible bases to span uncertainties of baryons in cosmic shear.

two baryon-contaminated mock data vectors based on the Illustris and Eagle scenarios. Throughout this work, when conducting a simulated analysis with a specific baryon-contaminated mock data vector, we avoid using this specific baryonic scenario as input to the construction of our baryon mitigation model. Further details of the simulated likelihood analyses are found in Section 2.3.

We derive the baryon-contaminated data vectors at a specific cosmology $\boldsymbol{p}_{\rm co}$ using the underlying hydrodynamical power spectrum defined as

$$P_\delta^{\rm hydro}(k, z \mid \boldsymbol{p}_{\rm co}) = \frac{P_\delta^{\rm hydro,sim}(k, z \mid \boldsymbol{p}_{\rm co,sim})}{P_\delta^{\rm DMO,sim}(k, z \mid \boldsymbol{p}_{\rm co,sim})} P_\delta^{\rm theory}(k, z \mid \boldsymbol{p}_{\rm co}). \quad (1)$$

The ratio terms $\frac{P_\delta^{\rm hydro,sim}(k,z \mid \boldsymbol{p}_{\rm co,sim})}{P_\delta^{\rm DMO,sim}(k,z \mid \boldsymbol{p}_{\rm co,sim})}$ are computed from interpolating the power spectrum ratio table constructed from various simulations snapshots from $z = 0 \sim 3.5$.[9] Some of the selected snapshots are visualized in Fig. 1.

When using equation (1), we implicitly assume that baryonic effects and cosmology are independent. That is, we fix the ratio of $P_\delta(k, z)$ for each baryonic scenario, while the cosmological dependence is propagated through the theoretical power spectrum $P_\delta^{\rm theory}(k, z \mid \boldsymbol{p}_{\rm co})$. The expression of $P_\delta^{\rm hydro}(k, z \mid \boldsymbol{p}_{\rm co})$ computed as in equation (1) is then passed into the COSMOLIKE package to derive the baryon-contaminated data vectors (Section 2.2).

The dependence of the baryonic suppression of the power spectrum with cosmological parameters is explored in detail in Schneider et al. (2020a). Based on the parametrization of their baryon correction model, the derived power spectrum ratios are largely independent of individual cosmological parameters, but they are related to the cosmic baryon fraction, $f_{\rm b} = \Omega_{\rm b}/\Omega_{\rm m}$ (see their Fig. 2); when varying $f_{\rm b}$ from 0.16 to 0.2, the power spectrum ratios are further suppressed by ∼5 per cent (10 per cent) for $k \leqslant 3\,h\,{\rm Mpc}^{-1}$ ($k \leq 10\,h\,{\rm Mpc}^{-1}$). Here we use $k \sim 3\,h\,{\rm Mpc}^{-1}$ as a reference, because this roughly corresponds to the effective scale to which we are sensitive given our small-scale cut at 2.5 arcmin, given the lensing kernel peak at $z \sim 0.5$ for DES Y1. Similarly, based on a fixed sets of BAHAMAS runs, but varying cosmologies from WMAP 2009 ($f_{\rm b} \sim 0.17$; Hinshaw et al. 2013), Planck 2013 ($f_{\rm b} \sim 0.15$; Planck Collaboration XVI 2014), to Planck 2015 ($f_{\rm b} \sim 0.16$; Planck Collaboration VIII 2016), van Daalen et al. (2020) showed that the power spectrum ratios vary $\lesssim$ 2 per cent (4 per cent) for $k \leqslant 3\,h\,{\rm Mpc}^{-1}$ ($k \leq 10\,h\,{\rm Mpc}^{-1}$). We note that the interaction between baryonic physics and cosmological parameters is a subdominent effect given the constraining power of DES Y1. According to Schneider et al. (2020a), ignoring the coupling of baryon suppression with $f_{\rm b}$ is a valid approximation even for future stage IV weak lensing surveys (see their fig. 10).

[9]The power spectra ratio data are available at https://github.com/hungjinh/baryon-power-spectra and from the van Daalen et al. (2020) data release.

2.2 Model

We use the COSMOLIKE package (Krause & Eifler 2017), one of the pipelines for DES cosmological inference, to perform the theoretical modelling of the 3 × 2pt data vectors. The linear DMO power spectrum is generated at each cosmology using CLASS (Blas, Lesgourgues & Tram 2011), with non-linear corrections derived from the Takahashi et al. (2012) version of HALOFIT. Throughout this work, we consider a flat ΛCDM cosmological model with six free parameters, $\boldsymbol{p}_{\rm co} = \{\Omega_{\rm m},\ A_{\rm s},\ \Omega_{\rm b},\ n_{\rm s},\ \Omega_\nu h^2,\ h\}$ in addition to the considered systematics parameters. The complete list of all parameters and their priors is given in Table 1.

Below we briefly summarize the theoretical modelling of the three types of two-point correlation functions and their associated systematic effects.

2.2.1 *Cosmic shear $\xi_\pm(\theta)$*

The real-space cosmic shear correlation function in tomographic bins i, j is modelled as

$$\xi_\pm^{ij}(\theta) = (1 + m^i)(1 + m^j)\,\frac{1}{2\pi}\int {\rm d}\ell\ \ell J_{0/4}(\ell\theta) C_{\gamma\gamma}^{ij}(\ell). \quad (2)$$

Here J_0 and J_4 are Bessel functions of the first kind. The m^i are multiplicative factors, one for each tomographic bin, that account for shear calibration bias (Heymans et al. 2006; Huterer et al. 2006). $C_{\gamma\gamma}^{ij}(\ell)$ is the detected shear–shear power spectrum, which contains the real lensing signal due to gravity (GG) as well as the contamination due to intrinsic alignment (II, GI, IG terms)

$$C_{\gamma\gamma}^{ij}(\ell) = C_{\rm GG}^{ij}(\ell) + C_{\rm II}^{ij}(\ell) + C_{\rm GI}^{ij}(\ell) + C_{\rm IG}^{ij}(\ell)\,. \quad (3)$$

Adopting the Limber approximation and the flat Universe assumption (these modelling assumptions are demonstrated to be sufficient for Y1, see Fang et al. 2020) the real lensing contribution can be computed as

$$C_{\rm GG}^{ij}(\ell) = \int_0^{\chi_{\rm h}} {\rm d}\chi_{\rm l}\,\frac{g^i(\chi_{\rm l}) g^j(\chi_{\rm l})}{\chi_{\rm l}^2} P_\delta\left(k = \frac{\ell}{\chi_{\rm l}}, \chi_{\rm l}\right), \quad (4)$$

where $\chi_{\rm l}$ is the comoving distance for the matter distribution (lens) along the line of sight, and $\chi_{\rm h}$ is the comoving horizon distance. The lensing kernel in the i-th tomographic interval is

$$g^i(\chi_{\rm l}) = \frac{3}{2}\frac{H_0^2\Omega_{\rm m}}{c^2}\frac{\chi_{\rm l}}{a(\chi_{\rm l})}\int_{\chi_{\rm l}}^{\chi_{\rm h}} {\rm d}\chi_{\rm s} n_{\rm s}^i(\chi_{\rm s})\frac{\chi_{\rm s} - \chi_{\rm l}}{\chi_{\rm s}}\,, \quad (5)$$

with $n_{\rm s}^i(\chi_{\rm s})$ being the probability density function (PDF) for the redshift distribution of source galaxies in tomographic bin i, defined such that $n_{\rm s}^i(\chi_{\rm s}){\rm d}\chi_{\rm s} = n_{\rm s}^i(z){\rm d}z$, which is normalized to unity.

For the intrinsic alignment (IA) contamination, we compute the intrinsic–intrinsic shape correlation due to the local tidal gravitational

field on pairs of source galaxies as,

$$C^{ij}_{\rm II}(\ell) = \int_0^{\chi_{\rm h}} {\rm d}\chi_{\rm s} \frac{n^i_{\rm s}(\chi_{\rm s}) n^j_{\rm s}(\chi_{\rm s})}{\chi^2_{\rm s}} P_{\rm II}\left(k = \frac{\ell}{\chi_{\rm s}}, \chi_{\rm s}\right). \tag{6}$$

The lensing shear-intrinsic shape correlations for pairs of galaxies where the foreground one is tidally torqued and the background one is sheared by the same gravitational field reads,

$$C^{ij}_{\rm GI}(\ell) + C^{ij}_{\rm IG}(\ell) = \int_0^{\chi_{\rm h}} {\rm d}\chi \frac{g^i(\chi) n^j_{\rm s}(\chi) + n^i_{\rm s}(\chi) g^j(\chi)}{\chi^2} P_{\rm GI}\left(k = \frac{\ell}{\chi}, \chi\right). \tag{7}$$

The $P_{\rm II}$ and $P_{\rm GI}$ are IA power spectra. Throughout the work, we adopt the commonly used non-linear alignment (NLA) model (Hirata & Seljak 2004) to mitigate IA uncertainties, i.e. assuming the amplitudes of IA power spectra are linearly related to the local density field:

$$\begin{aligned} P_{\rm II}(k, z) &= A^2(z) P_\delta(k, z) \\ P_{\rm GI}(k, z) &= A(z) P_\delta(k, z) \\ A(z) &= -A_{\rm IA} C_1 \frac{3 H_0^2 \Omega_{\rm m}}{8\pi G} D^{-1}(z) \left(\frac{1+z}{1+z_0}\right)^{\eta_{\rm IA}}. \end{aligned} \tag{8}$$

Here $D(z)$ is the linear growth factor; C_1 is the normalization constant being set at 5×10^{-14} $\rm M_\odot^{-1} h^{-2} Mpc^3$ (Brown et al. 2002); the pivot redshift z_0 is being set to 0.62. The nuisance parameters that go into the pipeline for IA marginalization are $A_{\rm IA}$ and $\eta_{\rm IA}$. For a more detailed IA analysis on DES Y1 data see Samuroff et al. (2018).

2.2.2 *Galaxy clustering*

The location of galaxies traces the underlying matter density field, yet with some unknown bias factor which depends on scales and redshift and on the tracer galaxy population. On large scales, under the simple scale-independent linear bias model, the theoretical prediction for the galaxy–galaxy autocorrelation function in tomographic bin i can be expressed as

$$\begin{aligned} w^i(\theta) &= \frac{1}{2\pi} \int {\rm d}\ell J_0(\ell\theta) C^{ii}_{\delta_{\rm g}\delta_{\rm g}}(\ell) \\ C^{ii}_{\delta_{\rm g}\delta_{\rm g}}(\ell) &= \left(b^i_{\rm g}\right)^2 \int_0^{\chi_{\rm h}} {\rm d}\chi_{\rm l} \frac{\left(n^i_{\rm l}(\chi_{\rm l})\right)^2}{\chi^2_{\rm l}} P_\delta\left(k = \frac{\ell}{\chi_{\rm l}}, \chi_{\rm l}\right), \end{aligned} \tag{9}$$

where $n^i_{\rm l}(\chi_{\rm l})$ is the PDF for the redshift distribution of lens galaxies, and $b^i_{\rm g}$ is the galaxy bias factor for each tomographic bin.

2.2.3 *Galaxy–galaxy lensing*

Galaxy–galaxy lensing, the cross-correlation between the position of lens galaxies in bin i and their surrounding matter density field traced by the shear of source galaxies in bin j, is modelled as

$$\gamma_{\rm t}^{ij}(\theta) = (1 + m^j) \frac{1}{2\pi} \int {\rm d}\ell J_2(\ell\theta) C^{ij}_{\delta_{\rm g}\gamma}(\ell), \tag{10}$$

where m^j again is the multiplicative shear bias; J_2 is the second-order Bessel function. Similarly, the $C^{ij}_{\delta_{\rm g}\gamma}(\ell)$ term has contributions from both pure lensing and IA effects,

$$C^{ij}_{\delta_{\rm g}\gamma}(\ell) = C^{ij}_{\delta_{\rm g}{\rm G}}(\ell) + C^{ij}_{\delta_{\rm g}{\rm I}}(\ell). \tag{11}$$

The lensing term reads

$$C^{ij}_{\delta_{\rm g}{\rm G}}(\ell) = b^i_{\rm g} \int_0^{\chi_{\rm h}} {\rm d}\chi_{\rm l} \frac{n^i_{\rm l}(\chi_{\rm l}) g^j(\chi_{\rm l})}{\chi^2_{\rm l}} P_\delta\left(k = \frac{\ell}{\chi_{\rm l}}, \chi_{\rm l}\right), \tag{12}$$

and the IA term is expressed as

$$C^{ij}_{\delta_{\rm g}{\rm I}}(\ell) = b^i_{\rm g} \int_0^{\chi_{\rm h}} {\rm d}\chi \frac{n^i_{\rm l}(\chi) n^j_{\rm s}(\chi)}{\chi^2} P_{\rm GI}\left(k = \frac{\ell}{\chi}, \chi\right), \tag{13}$$

with the IA power spectrum $P_{\rm GI}$ being defined in equation (8).

Finally, throughout this work, the uncertainty in the photometric redshifts is modelled as a constant shift of the initial redshift probability distribution function $n^i_{\rm pz}(z)$, for both source and lens galaxies, in each tomographic bin.

$$n^i_{\rm s}(z) = n^i_{\rm s,pz}\left(z - \Delta z^i_{\rm s}\right) \quad ; \quad n^i_{\rm l}(z) = n^i_{\rm l,pz}\left(z - \Delta z^i_{\rm l}\right). \tag{14}$$

2.3 PC decomposition to model baryonic effects

We adopt the principal component (PC) decomposition technique to model baryonic effects for small-scale cosmic shear (Eifler et al. 2015). The basic idea of this technique is to perform principal component analysis (PCA) on the difference of the theoretical model vectors (the 3×2pt vectors for this work) between hydrodynamical and DMO simulations, for several baryonic scenarios. To construct the baryon-contaminated cosmic shear correlation functions, we use equation (1) to derive the underlying baryonic power spectra that go into integration (see Section 2.1.3).

The resulting dominant PC modes then serve as a flexible basis set to account for possible baryonic effects in both spatial and redshift dimensions via the angular bins and tomographic information. In H19, we validate this method assuming an LSST-like cosmic shear experiment. We further improve the efficiency of this method by imposing a covariance-driven weighting factor when performing PCA, which is referred to as method C in H19. Below we briefly summarize the formalism of this method.

Let $\boldsymbol{M}$ be a DMO-based theoretical 3×2pt model vector, and $\boldsymbol{B}_x$ be a model vector contaminated with baryonic scenario x, computed by replacing the matter power spectrum via equation (1) (see Section 2.1.3 for detail). We first build a difference matrix $\boldsymbol{\Delta}$

$$\boldsymbol{\Delta} = \left[\boldsymbol{B}_1 - \boldsymbol{M}\ \boldsymbol{B}_2 - \boldsymbol{M}\ \ldots\ \boldsymbol{B}_{N_{\rm sim}} - \boldsymbol{M}\right]_{N_{\rm data} \times N_{\rm sim}}. \tag{15}$$

Each column of $\boldsymbol{\Delta}$ is a difference vector, $\boldsymbol{B}_x - \boldsymbol{M}$, with 630 elements (Section 2.1.1), computed with the cosmology and the nuisance parameters being set to the fiducial values listed in Table 1.

Next we use the Cholesky decomposition on the data vector covariance matrix, to find the square root of the covariance

$$\mathbf{C} = \mathbf{L}\mathbf{L}^{\rm t}. \tag{16}$$

We use $\mathbf{L}^{-1}$ to build a noise-weighted difference matrix $\boldsymbol{\Delta}_{\rm ch}$, and apply singular value decomposition (SVD) to $\boldsymbol{\Delta}_{\rm ch}$

$$\begin{aligned} \boldsymbol{\Delta}_{\rm ch} &= \mathbf{L}^{-1} \boldsymbol{\Delta} \\ &= \mathbf{L}^{-1} \left[\boldsymbol{B}_1 - \boldsymbol{M}\ \boldsymbol{B}_2 - \boldsymbol{M}\ \ldots\ \boldsymbol{B}_{N_{\rm sim}} - \boldsymbol{M}\right]_{N_{\rm data} \times N_{\rm sim}} \\ &= \mathbf{U}_{\rm ch}\ \boldsymbol{\Sigma}_{\rm ch}\ \mathbf{V}^{\rm t}_{\rm ch}, \end{aligned} \tag{17}$$

where $\mathbf{U}_{\rm ch}$ and $\mathbf{V}_{\rm ch}$ are square unitary matrices with dimensions of $N_{\rm data} \times N_{\rm data}$ and $N_{\rm sim} \times N_{\rm sim}$, respectively. $\boldsymbol{\Sigma}_{\rm ch}$ is a diagonal matrix with the singular values populating the diagonal in descending order.

The first $N_{\rm sim}$ columns of the $\mathbf{U}_{\rm ch}$ matrix form a set of PC bases, $\mathbf{v}_{{\rm PC},i}$, that can be used to fully span the baryonic features of our training simulations. For a given baryonic scenario x, we have

$$\mathbf{L}^{-1}(\boldsymbol{B}_x - \boldsymbol{M}) = \sum_{i=1}^{N_{\rm sim}} Q_i\ \mathbf{v}_{{\rm PC},i}. \tag{18}$$

With the derived PCs, we can generate a baryonic model that utilizes PC amplitudes Q_i to simulate possible baryonic behaviours.

$$\boldsymbol{M}_{\rm bary}(\boldsymbol{p}_{\rm co}, \boldsymbol{p}_{\rm nu}, \boldsymbol{Q}) = \boldsymbol{M}(\boldsymbol{p}_{\rm co}, \boldsymbol{p}_{\rm nu}) + \sum_{i=1}^{n} Q_i \, \mathbf{L} \cdot \mathbf{v}_{{\rm PC},i} \ . \tag{19}$$

Here n specifies the number of PC amplitudes/PC modes used to model the baryonic effect, and $n \leq N_{\rm sim}$. The operation of $\mathbf{L} \cdot \mathbf{v}_{{\rm PC},i}$ transforms the PC mode back to the same basis as $\boldsymbol{M}$.

Note that although we pass the full 3×2pt vector in equation (17) to perform PCA, the deviations from the DMO scenario are extremely small for the galaxy–galaxy lensing and galaxy clustering parts because of their conservative scale cuts. Therefore, the PCs mostly account for baryonic effects in small-scale data points of cosmic shear (see Fig. A2 for the fractional change of model vector when varying Q_1).

2.3.1 *Input hydrodynamical scenarios for PC construction*

We will use the Illustris and the Eagle scenarios as the conservative and optimistic validation scenarios for our PCA baryon mitigation model.

As mentioned before, we exclude the considered scenario in the baryon PC basis set, hence we are building two PC bases for this exercise, one excluding Eagle and the second one excluding Illustris.

The first PC set is constructed with 10 hydrodynamical scenarios: MB2, Horizon-AGN, TNG100, Eagle, three variants of cosmo-OWLS, and three variants of BAHAMAS scenarios with different AGN feedback strength. We will use this basis set to mitigate baryonic effects for our Illustris and DMO mock data vectors (see Section 2.1.3), and for the real DES Y1 observational data vector.

The second PC basis set is constructed with the same scenarios as the first, with the Eagle scenario being excluded. When performing our analyses on the Eagle mock data vector, we will use the second PC set as bases to conduct baryon mitigation. The reasoning for this design can be understood in equation (18). If using the first PC set to perform marginalization on Eagle, the first PC set is guaranteed to be able to describe Eagle by construction.

Fig. 2 provides a visualization of $\mathbf{L} \cdot \mathbf{v}_{{\rm PC},i}$ in projection on the $\xi_\pm$ observables in the cross tomographic bin (2,3), for our two sets of PC bases. As shown, these two sets of PC modes turn out to be quite similar.

2.4 Likelihood analysis

We infer the posterior probability distribution of cosmological ($\boldsymbol{p}_{\rm co}$) and nuisance parameters ($\boldsymbol{p}_{\rm nu}$) via Bayes' theorem:

$$P(\boldsymbol{p}_{\rm co}, \boldsymbol{p}_{\rm nu}|\boldsymbol{D}) \propto L(\boldsymbol{D}|\boldsymbol{p}_{\rm co}, \boldsymbol{p}_{\rm nu}) P_{\rm prior}(\boldsymbol{p}_{\rm co}, \boldsymbol{p}_{\rm nu}) \ , \tag{20}$$

with the prior probability distribution for each of the parameter defined in Table 1.

The priors for our baseline analysis are chosen to be mostly the same as DES Y1 (Abbott et al. 2018), with an exception of the upper limit of the neutrino mass prior. We now discuss the priors in more detail.

2.4.1 *Prior for neutrino mass*

Instead of applying a non-informative upper limit on the sum of neutrino masses $\Sigma m_\nu < 0.94\,{\rm eV}\,c^{-2}$ (i.e. $\Omega_\nu h^2 < 0.01$) as in DES Y1, we adopt an upper limit of $\Sigma m_\nu < 0.12\ {\rm eV}\,c^{-2}$ ($\Omega_\nu h^2 < 0.0013$) as our baseline analysis. This upper limit is based on the latest 95 per cent constraint from *Planck* TT, TE, EE+lowE+lensing+BAO (Planck Collaboration VI 2018) and has the advantage that it reduces biases in the 1D projected posterior probabilities of the relevant cosmological parameters.

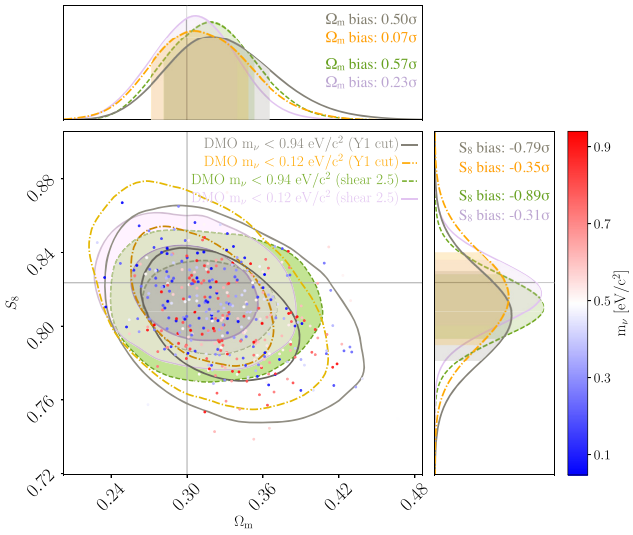

Figure 3. Cosmological parameter constraints from simulated likelihood analyses subjected to different choices of neutrino mass priors. All chains shown here are based on the DMO mock 3×2pt analyses. The grey (orange) contours indicate the case with fiducial Y1 scale cuts, and with wide (narrow) neutrino prior applied. The shaded green (pink) contours are the case when extending cosmic shear to 2.5 arcmin in the 3×2pt mock data (without performing baryon marginalization), subjected to the Y1 fiducial (narrow) neutrino prior. Here, and in all 2D posterior plots below, the contours depict the 68 per cent and 95 per cent confidence levels. The parameter biases of $\Omega_{\rm m}$ and S_8 decrease when narrowing the neutrino prior. The coloured dots are randomly selected samples in the wide neutrino prior chain (orange curves) with the neutrino mass coloured as indicated in the sidebar. Higher neutrino mass tends to suppress the clustering amplitude of matter. *The posterior of* $\Omega_{\rm m}$ is thus biased high to compensate for that.

As shown in Fig. 3, for the simulated likelihood analysis with DES Y1 scale cuts, the wide Y1 neutrino prior leads to a $\sim 0.8\sigma$ level bias in S_8 (grey contour); while the case with an informative neutrino prior only has a 0.35σ bias in S_8. The bias is caused by the asymmetric coverage of the neutrino prior around the fiducial neutrino value, as discussed in Krause et al. (2017). Since DES Y1 data do not have significant constraining power on neutrino masses, marginalizing over neutrino mass preferentially allows many scenarios with increased neutrino mass, which leads to a net suppression in structure growth. The $\Omega_{\rm m}$ posterior is then biased high to compensate for that.

This neutrino prior-induced bias becomes more significant when including small-scale data in the analysis, due to the smaller uncertainties (grey versus green contours in Fig. 3) and due to the fact that small-scale data are more sensitive to the neutrino mass. We thus place a narrower limit on the neutrino mass prior as our baseline analysis, and the S_8 bias is reduced to $\sim 0.3\sigma$ (shaded pink contours). For the purpose of comparison, we will also present and discuss the result with the original Y1 wide neutrino prior.

2.4.2 *Priors for baryonic parameters*

The theoretical PC amplitude Q_i for each hydrodynamical scenario x can be computed by taking the inner product of the weighted

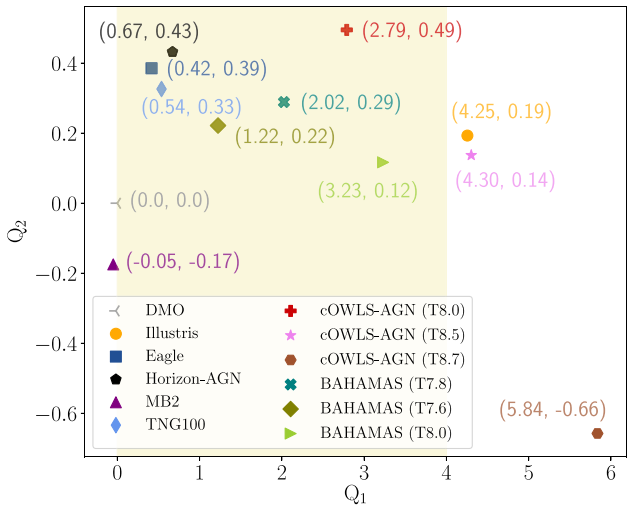

Figure 4. The expected PC amplitudes Q_1, Q_2 for baryonic scenarios considered in this work, with PCs being constructed using 10 hydrodynamical simulations as detailed in Section 2.3.1. Our choice of priors for $Q_{1,2}$ are based on the range of values from these baryonic scenarios. The yellow band highlights the range of informative prior on Q_1, Flat(0,4), which is bounded by the Illustris scenario. As shown in Fig. 2 on the features of PC modes, the larger Q_1 value indicates stronger suppression of matter clustering at small scales. The Q_2 parameter further provides a higher order correction.

difference vector, $\mathbf{L}^{-1}(\boldsymbol{B}_x - \boldsymbol{M})$, with the PC mode PC_i (see equation 18). Fig. 4 presents the expected $Q_{1,2}$ values for all hydrodynamical scenarios considered in this work.

An increase in Q_1 mostly controls the amount of suppression on small scales, whereas higher order PC amplitudes $Q_{i\geq 2}$ provide corrections on baryonic effects that can also impact larger scales (see Fig. 2).

We adopt two choices of priors for the baryonic parameters.

(i) baseline: $Q_1 \in \mathrm{Flat}(-3, 12)$; $Q_2 \in \mathrm{Flat}(-2.5, 2.5)$
(ii) informative: $Q_1 \in \mathrm{Flat}(0, 4)$; $Q_2 \in \mathrm{Flat}(-2.5, 2.5)$

The baseline priors are extremely conservative and allow for the data to entirely self-calibrate the baryonic effects. Looking at Fig. 4 we see that they are significantly larger than the spread of $Q_{1,2}$ for all hydrodynamical scenarios.

The informative prior of Q_1 is highlighted in the yellow band of Fig. 4, and we consider this prior range to be well-motivated if one considers including a minimal amount of external information from the simulation literatures in our analysis. Specifically, Haider et al. (2016) found that the radio-mode AGN feedback in Illustris is too strong such that too much gas is heated and ejected, leading to insufficient baryons in galaxy groups compared with observations. Also in Le Brun et al. (2014), the cosmo-OWLS T8.5 and T8.7 scenarios are a poor match to several X-ray observables. It is thus reasonable to use the Illustris scenario as an upper bound on the level of feedback strength that our Universe could possibly reach and to adopt a corresponding prior in our analysis. As we will see later, this informative, but well-motivated prior, will increase the amount of information we gain on cosmology by adding small-scale cosmic shear data in DES Y1.

For our likelihood analyses, we adopt a Gaussian likelihood:

$$L(\boldsymbol{D}|\boldsymbol{p}_{\rm co}, \boldsymbol{p}_{\rm nu}) \propto \exp\left(-\frac{1}{2}\underbrace{\left[(\boldsymbol{D}-\boldsymbol{M})^t\,\mathbf{C}^{-1}\,(\boldsymbol{D}-\boldsymbol{M})\right]}_{\chi^2(\boldsymbol{p}_{\rm co},\boldsymbol{p}_{\rm nu})}\right). \quad (21)$$

As discussed in Lin et al. (2019), the impact of non-Gaussianity in the likelihood is estimated to be negligible in current and future cosmic shear surveys.

The 3×2pt covariance matrix $\mathbf{C}$ is computed using the COSMOLIKE package (Krause & Eifler 2017), which calculates the relevant four-point functions in the halo model. The analytic form of the covariance matrix and relevant validation for DES Y1 is detailed in Krause et al. (2017), with updates provided in Troxel et al. (2018b) to address the effect of survey geometry and the uncertainty in the multiplicative shear bias calibration. For simplicity, we do not consider potential baryonic effects when computing the covariance matrix. As discussed in previous works, neglecting baryonic effects in the covariance matrix has little impact on the cosmological inference for stage IV weak lensing surveys Schneider et al. (2020a), Barreira et al. (2019), and should be negligible for DES.

We use the `emcee` package (Foreman-Mackey et al. 2013), which relies on the affine-invariant ensemble sampling algorithm (Goodman et al. 2010), to sample the parameter space. We run MCMC (Markov Chain Monte Carlo) chains to 2.5 million steps, and then discard the first 1.25 million steps as burn-in. We have visually checked the convergence of MCMC chains by ensuring that the 1D and 2D posterior distributions for all parameters are consistent with the results of a chain with 5 million steps out to 3σ confidence intervals.

2.5 Blinding strategy

Our blinding strategy aims to shield against 'confirmation bias', i.e. stopping the search for new systematics or better parametrizations of existing systematics when the result matches the expectation. There are differences between our analysis and the DES Y1 analysis choices described in Krause et al. (2017), and these analysis differences will drive those in the respective blinding strategies. In particular, we include small scales in cosmic shear (down to 2.5 arcmin), we add a corresponding parametrization for baryonic physics uncertainties, and we use a different prior for the neutrino mass parameter. Beyond these differences, we follow the Krause et al. (2017) choices; in particular, we do not reassess scale cuts for galaxy–galaxy lensing and galaxy clustering, or other model parametrizations and priors. This is justified given that our constraining power is very similar to that of Abbott et al. (2018), and even when we use informative priors on baryonic physics we expect a 20 per cent information increase at most (see Section 3.2). Based on these considerations, our blinding strategy proceeds as follows:

(i) We develop our pipeline completely independently of the data vector, i.e. we run 100+ simulated likelihood analyses to stress test our pipeline. We use different data vectors, different prior settings, and different modelling settings until we converge to the setup described in Table 1. We describe this process in Section 2.4 and results in Section 3.

(ii) Our pipeline is a modified version of the DES Y1 COSMOLIKE pipeline; we performed a comparison with the latest COSMOLIKE version (which has undergone testing and validation for DES Y3) and have reached an excellent agreement at the level of $\Delta\chi^2 = 0.0005$ and $\Delta\chi^2 = 0.0006$ for model vectors with the original scale cuts used in Abbott et al. (2018) and with the new scale cuts used in this paper, respectively. The residual uncertainties are due to small modifications in the interpolation routines that were incorporated between Y1 and Y3. The version used in this work is tagged as 'Huang2020' in the 'cosmolike_core' github repository of the COSMOLIKE github organization.

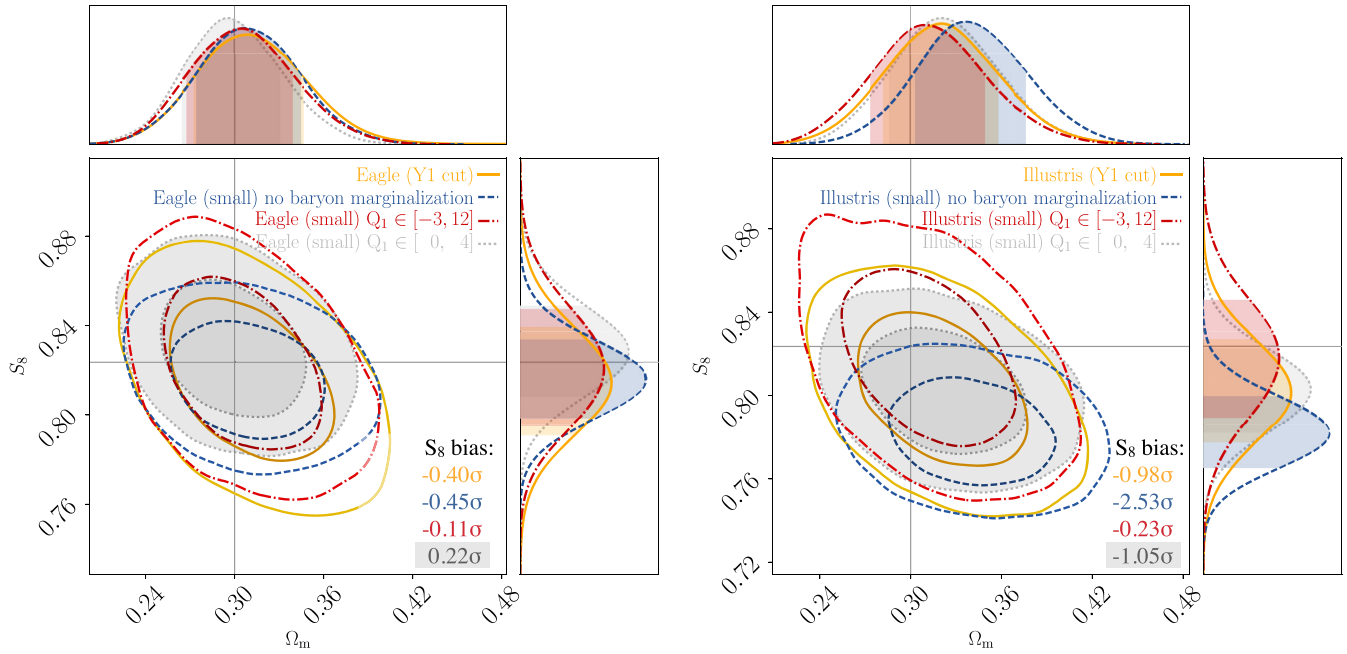

Figure 5. Results with the analysis setup where we fit the 3×2pt observables while marginalizing over baryonic physics on two hydrodynamical simulations. The two panels show posterior constraints on Ω_m and S_8 for the Eagle (left-hand panel) and the Illustris (right-hand panel) mock data. The yellow contours indicate the result when the DES Y1 scale cut is applied. The blue (red) contours show the result when extending cosmic shear to 2.5 arcmin, but without (with) marginalization on baryonic parameter Q_1. The analysis result when adopting the informative Q_1 prior is indicated in grey shaded contours. The marginalized 1D constraint in S_8 with 1σ error is spelled out in the lower right corners of the plots.

(iii) We described all our pipeline tests and the code comparison to an internal review panel within the DES collaboration and only replaced the simulated data vector with the actual data after their sign-off.

The data constraining results presented in Section 5 and Section 4 are unaltered post unblinding.

3 LIKELIHOOD SIMULATION RESULTS

In this section, we present our simulated likelihood analysis results for the three mock data vectors of baryonic scenarios, DMO, Eagle, and Illustris, in order to design and understand the expected performances of our baryon mitigation pipeline. DMO is the best-case scenario for which we know in advance that the resulting cosmological inference should not be biased, regardless of whether baryon mitigation is performed. With its strong feedback, the Illustris simulation serves as a conservative scenario in our pipeline validation; such strong feedback is largely ruled out by observations already (Haider et al. 2016). The Eagle scenario has significantly weaker feedback, so its deviation from DMO is relatively small and it serves as an optimistic scenario in our pipeline validation.

As an overview, in Fig. 5, we show the posterior distributions of Ω_m and S_8 with the input mock 3 × 2pt data from the Eagle (left-hand panel) and Illustris (right-hand panel) scenarios. We compare the fiducial DMO case (filled grey contours) with

(i) applying DES Y1 scale cuts (yellow contours),
(ii) extending cosmic shear to 2.5 arcmin but without introducing an extra parameter to marginalize over baryonic physics (blue contours),
(iii) same as (ii), but marginalizing over Q_1 with our baseline prior Flat(−3, 12) (red contours),
(iv) same as (iii), but applying an informative prior Flat(0, 4) on Q_1 (grey shaded contours).

Below we will investigate the posterior distributions on these simulated likelihood analyses (shown in Fig. 5), to understand the potential outcomes when applying our pipeline on real data.

3.1 Number of PC modes to be marginalized over given DES Y1 constraining power

To determine how many PC modes are needed in equation (19) to account for baryons when pushing cosmic shear to 2.5 arcmin given DES Y1 statistical power, we increase the available degrees of freedom by increasing the number of PC amplitudes Q_i when running likelihood simulations and track the resulting posterior distributions.

3.1.1 The residual bias after marginalization

Fig. 6 summarizes the marginalized 1D S_8 posterior constraints for our likelihood simulations (as shown in Fig. 5 for the case of the Eagle and Illustris scenarios).

We use the DMO results as a baseline for understanding the level of parameter projection effects, i.e. the parameter biases as revealed in the marginalized posterior constraints. Parameter degeneracies in the high-dimensional space may lead the 1D and 2D projected posteriors to peak at biased positions, and for parameters where the data are not sufficiently constraining the posterior can peak at biased values due to prior volume effects (e.g. the neutrino prior issue discussed in Section 2.4.1). As indicated by the yellow square markers, we observe that the projection effects would cause $\approx 0.3 \sim 0.5\sigma$ biases in the S_8 constraints under our baseline setting (see Table 1), which we should keep in mind when interpreting tensions between different experiments using distances in projected parameter spaces.

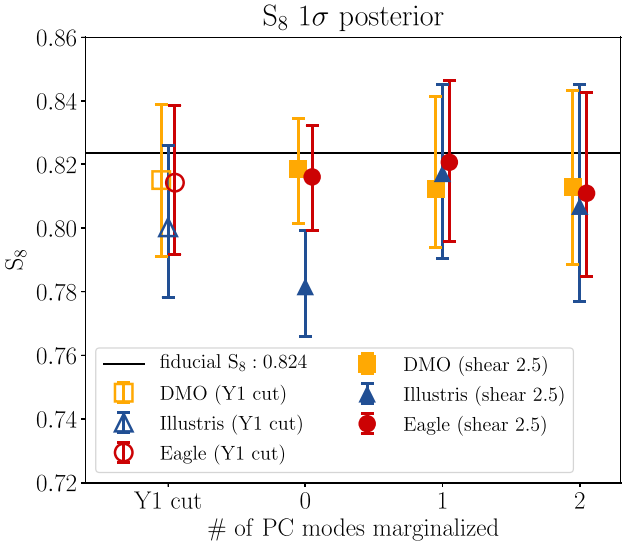

Figure 6. The marginalized 1D S_8 posterior constraints for our baseline likelihood simulations with baryonic scenarios of DMO, Eagle, and Illustris. Each marker's centre, lower, and upper error bars indicate the median, the 16th and the 84th percentiles of marginalized 1D posteriors. The open markers are results of 3 × 2pt mock data vectors subjected to the original DES Y1 scale cuts, while the filled markers are results when extending the cosmic shear data points to 2.5 arcmin (Section 2.1), with different choices for the number of marginalized PC amplitudes Q_i to account for baryonic effects. Marginalizing over 1 PC mode is sufficient to account for baryonic effects to within $\sim 0.2\sigma$ under the statistical power of DES Y1, for all baryonic scenarios considered here.

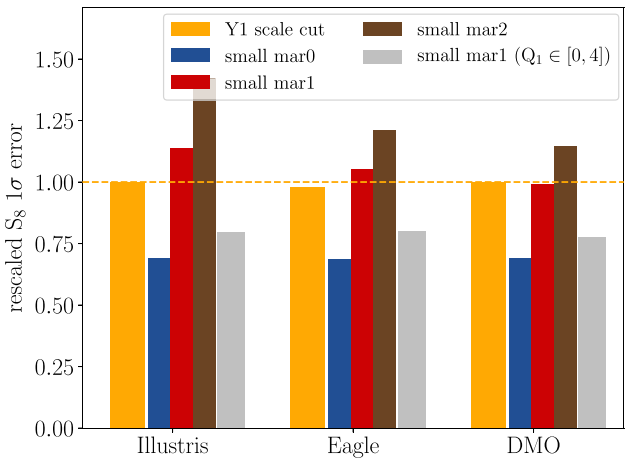

Figure 7. The rescaled S_8 1σ error in the simulated likelihood analyses: $\sigma/\sigma_{\rm Y1\,cut\,DMO}$. The horizontal yellow dashed line indicates the condition when $\sigma = \sigma_{\rm Y1\,cut\,DMO}$. Starting from left to right, the first/second/third groups indicate results when using the Illustris/Eagle/DMO scenarios as mock data in our likelihood simulations. The yellow bars are for S_8 errors derived with Y1 scale cuts applied. The blue/red/brown bars are the results with the cosmic shear data points extended to 2.5 arcmin, and with 0/1/2 PC mode(s) being marginalized. The grey bars indicate the results when marginalizing over Q_1 with an informative prior range. For the baseline analysis setting with small-scale cosmic shear included, marginalizing over 1 PC mode leads to similar constraining power in S_8 compared with the result with Y1 scale cuts being applied. When adopting an informative prior on Q_1, a $\sim$ 20 per cent improvement in S_8 is expected.

When performing analyses with the Y1 scale cut (open markers) without marginalizing over baryonic physics, we find a residual $\sim 0.9\sigma$ bias in S_8 for the Illustris scenario. This is because the Y1 cosmic shear scale cut is determined based on the cOWLS-AGN $\Delta T_{\rm heat} = 10^8$ scenario, which is less intense compared with the feedback effect of Illustris (see Fig. 1). When including small-scale cosmic shear data points in the 3 × 2pt analyses (filled markers), for weaker baryonic scenarios like Eagle, we find that even without marginalization the S_8 bias can still be within 0.5σ. Using a strong feedback scenario like Illustris as the most pessimistic limit, we conclude that marginalizing over a single PC mode would be sufficient to account for baryonic effects to within $\sim 0.2\sigma$, which is well within the referential bias level set from the DMO case.

3.1.2 The degradation on parameter constraints after marginalization

Small-scale cosmic shear data points provide additional cosmological information, but some of the information will be lost after accounting for uncertainties in baryonic physics. Here, we explore the expected degradation on parameter constraints in DES Y1 within the PCA framework, subject to our choices on the number of marginalization parameters for baryons.

Fig. 7 shows the rescaled S_8 1σ error for our likelihood analyses (as shown in Fig. 5 for the cases of the Eagle and Illustris scenarios). Starting from left to right, the first/second/third group is for the Illustris/Eagle/DMO mock data vectors when running likelihood simulation. The yellow bars are for S_8 errors derived with Y1 scale cuts applied. The blue/red/brown bars are the results with the cosmic shear data points extended to 2.5 arcmin, and with 0/1/2 PC mode(s) being marginalized. This figure confirms our expectation that after marginalizing over one PC mode (red bars), the resulting S_8 constraint should be similar to the result with conservative scale cuts being applied (yellow bars). Marginalizing over two PC modes (brown bars) should lead to 20 per cent$\sim$30 per cent larger errors in S_8, depending on the baryonic scenarios.

In conclusion, we do not expect to gain extra cosmological information from small-scale cosmic shear data points when using our wide baseline prior to account for baryons. The same conclusion can be inferred from the 2D posterior distributions in the $\Omega_{\rm m}$–S_8 plane presented in Fig. 5 for the analyses using the Eagle and the Illustris scenarios as mock data.

3.2 Information gained with informative prior on baryonic physics

Next we explore the improvement in the constraints on cosmological parameters when adopting our well-motivated, informative Q_1 prior which limits the allowed range of baryonic uncertainties to exclude feedback strength at the Illustris level (Section 2.4.2). Thus, when adopting our informative prior, we do not expect the allowed degrees of freedom to fully mitigate Illustris or other baryonic scenarios with more intense AGN feedback.

As shown in the grey bars of Fig. 7, we expect to have about 20 per cent improvement in the marginalized 1D S_8 constraint when using the informative $Q_1 \in (0,4)$ prior, compared with adopting Y1 scale cuts (yellow bars). Fig. 5 also provides a visualization of the relative improvements in terms of 2D posterior distributions (grey shaded versus yellow contours).

3.3 Expected constraints on baryonic parameters

Next we present the expected constraints on the baryon parameters (PC amplitudes) for our baseline pipeline setting and

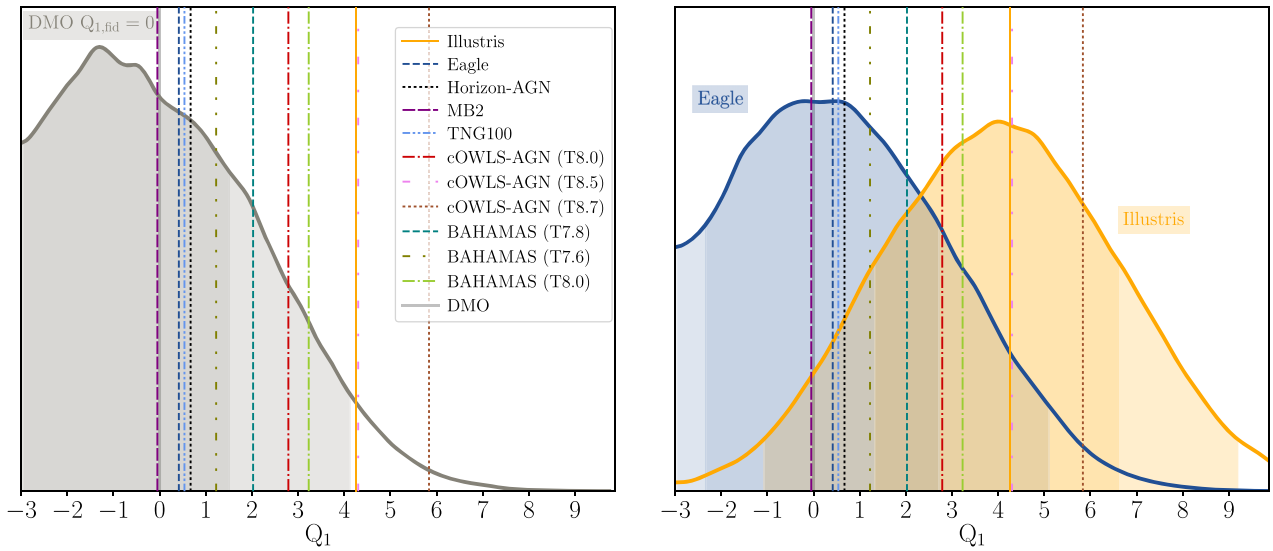

Figure 8. Posterior distributions of Q_1 under our baseline setting for mock data vectors of DMO (left-hand panel), and Illustris and Eagle scenarios (right-hand panel). We show the case when cosmic shear is extended to 2.5 arcmin, and only varying Q_1 to account for baryons. Given the constraining power of Y1, when the baryonic feedback of the input mock data is weak (at the level of Eagle), we expect to exclude some of the extreme baryonic scenarios at the $\sim 2\sigma$ confidence intervals.

discuss the potential parameter projection effects on their posterior distributions.

Fig. 8 shows the Q_1 posterior distributions for the DMO (left-hand panel) and Eagle and Illustris (right-hand panel) mock data vectors, with cosmic shear data down to 2.5 arcmin and with only Q_1 being marginalized over an un-informative prior. The theoretical values of Q_1 for various baryonic scenarios are computed relying on the relation of equation (18), as detailed in Section 2.4.2.

The DMO mock data are created with $Q_1 = 0$, and therefore could be used to estimate the level of projection effects on Q_1. We see that the marginalized 1D peak of Q_1 has a $\sim 0.5\sigma$ shift from its fiducial value. This happens due to parameter degeneracies between Q_1 and other cosmological and systematics parameters. As discussed in Section 3.1.1, we have seen that biases of the order of $\sim 0.3 \sim 0.5\sigma$ are expected (see the yellow square markers in Fig. 6) in the marginalized 1D S_8 constraints for the case of DMO. We explore the topic of parameter degeneracies in more detail in Appendix A.

Note that the projection effect in Q_1 is less apparent for the cases of Eagle ($\sim 0.01\sigma$) and Illustris ($\sim 0.1\sigma$), as shown in the right-hand panel of Fig. 8. This is because the mock data vectors of Eagle and Illustris have extra baryonic features hidden in the higher order PC modes. This residual 'noise' of baryonic physics, which cannot be accounted for using only Q_1, is pushing the Q_1 posterior closer to the expected theoretical value.

Regarding the constraining power on baryonic physics shown in Fig. 8, we find that the DES Y1 type constraint can exclude baryonic scenarios that are different from the input fiducial scenario by $\sim 2\sigma$. For example, when the input fiducial baryonic scenario is weak like Eagle (blue curve in the right-hand panel), the cosmoOWLS-AGN with the minimum heating temperature at $10^{8.7}$ K (the strongest feedback baryonic scenario in the pool) can be excluded at the 2σ confidence level, given Y1 statistical power. When the input mock baryonic scenario is Illustris (yellow curve in the right-hand panel), all other baryonic scenarios are covered within the 2σ posterior region of Q_1.

4 COSMOLOGY CONSTRAINTS FROM DES Y1 DATA

This section presents the main cosmology results when applying our pipeline to DES Y1 data.

Fig. 9 presents a summary of the 68 per cent confidence intervals on the constraints of $\Omega_{\rm m}$, S_8, and σ_8 for all the analyses we have run. As a high-level summary, we start by presenting the DES Y1-only constraints of the baseline setting with Y1 scale cuts applied (top orange), and with the cosmic shear data extended to 2.5 arcmin but without performing baryon marginalization (green). With baryonic effects being properly marginalized through an non-informative prior on Q_1, we find that almost no information is gained with the inclusion of small-scale cosmic shear data points (dark blue) compared to the case with conservative Y1 scale cuts. When using our informative prior, we find an improvement in cosmological constraints from the inclusion of small-scale cosmic shear (grey). Finally, we also present a result with cosmic shear being extended down to 2.5 arcmin but with an adoption of the non-informative prior on neutrino mass (purple). We then combine the DES Y1 constraints with external data sets of Planck and baryon acoustic oscillation (BAO) constraints (darker orange and light blue), and explore the combined results (red and yellow). We will discuss these results in detail below.

4.1 Baseline constraints

Our baseline setting mostly follows the fiducial DES Y1 3 × 2pt key paper (Abbott et al. 2018), except for adopting a informative neutrino prior of $m_\nu < 0.12$ eV c^{-2} (see Section 2.4).

Fig. 10 shows the best-fitting theoretical models on top of the observed cosmic shear correlation function. The yellow lines show the fits when the original Y1 scale cut is applied. With the discarded small-scale data points added in the analyses, the yellow-green contours are the result without performing baryon marginalization; the blue lines show the result when the first PC amplitude is used to marginalize over uncertainties in baryonic effects. In Table 2, we provide χ^2 analyses on the derived best-fitting models. For the case

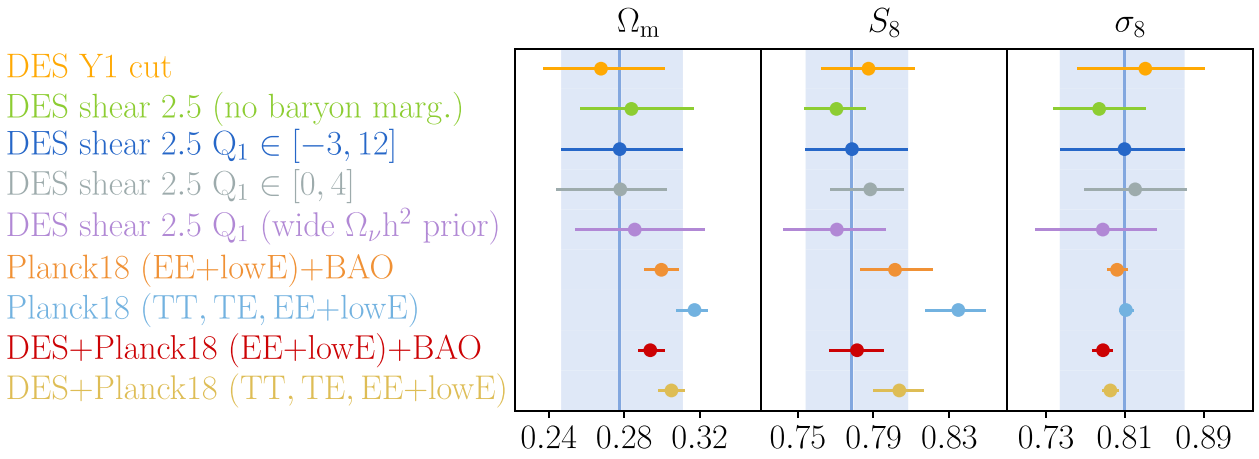

Figure 9. Summary of the marginal 1D peak values of cosmological parameters in ΛCDM from DES Y1 data. The 68 per cent confidence levels are shown as horizontal bars. The first row is the baseline DES 3 × 2pt data analysis with the adoption of the conservative Y1 scale cuts, and with the informative neutrino mass prior being applied. Rows 2–4 present the DES results with cosmic shear extended to 2.5 arcmin, but without taking baryonic uncertainty into account (yellow-green), using the Q_1 parameter to marginalize over the effect of baryons with an uninformative Q_1 prior (blue), and with an informative Q_1 prior (grey). The 5th row (purple) shows the result using the same setup as the second row, but adopting wide neutrino prior as in the original DES Y1 analysis. Two likelihood chains from the Planck DR18 results (Planck Collaboration VI 2018) are presented for comparison: the CMB polarization auto power spectra combined with BAO (orange), and the joint CMB temperature and polarization auto- and cross-power spectra (light blue). The last two rows are the results of the small-scale extended DES data (as shown in third row) when adopting informative cosmological priors from Planck EE+BAO (red), and from Planck TT,TE,EE (yellow).

of Y1 scale cut (first column), the reduced χ^2 derived in this work is consistent with the fiducial DES Y1 key paper (as discussed in the Appendix C of Abbott et al. 2018). After including the extra 175 small-scale data points of cosmic shear, but without introducing any new parameters in the modelling procedure, the reduced χ^2 value remains low (second column of Table 2).[10] This suggests that the baryonic features in the power spectrum on the scales to which these data are sensitive are weak enough that, within the Y1 error, the DMO calibrated theoretical model still provides a valid description of the data. Adding an extra degree of freedom to account for the potential baryonic effect at small scales does not reduce the χ^2 value any further.

The posterior cosmological parameter distributions from our baseline analyses are presented in Fig. 11. The yellow contours show the constraint when the original Y1 scale cuts are applied. The derived marginalized 1D peak constraints are

$$\Omega_m = 0.268^{+0.034}_{-0.031} \quad S_8 = 0.787^{+0.024}_{-0.025} \quad \sigma_8 = 0.831^{+0.060}_{-0.069}\,, \tag{22}$$

which is consistent with the Y1 key paper result. The minor difference in the constraints is caused by the following three factors: neutrino prior difference (which would induce a factor of $\sim 0.5\sigma$ shift as shown in Fig. 3), sampler difference (`emcee` versus `MultiNest` – Feroz, Hobson & Bridges 2009) and theory code uncertainty (COSMOLIKE versus COSMOSIS – Zuntz et al. 2015). As discussed in fig. 17 of Abbott et al. (2018), the latter two differences together contribute at about 0.2σ level. Given the Y1 constraining power, these uncertainties would not change the conclusion of this paper, but further investigation and management of their error budgets will become a necessity for Stage IV cosmological analyses.

[10]Note that for the small-scale data analyses, the χ^2 value goes from 674 to 675 when an additional degree of freedom is added to perform baryon marginalization. This could happen due to the stochastic MCMC sampling in high dimensional parameter space, and that the likelihood surface is not a simply smooth function but noisy. So when comparing the reduced χ^2 values, their error bars ($\sqrt{2/d.o.f}$) are important.

The blue contours in Fig. 11 show the cosmological constraints when introducing additional information from small-scale cosmic shear in the original Y1 3 × 2pt analysis while properly marginalizing over the uncertainties of baryons at these scales. The derived marginalized 1D posterior peaks of the main cosmological parameters are

$$\begin{aligned} \Omega_m &= 0.278^{+0.034}_{-0.031} \\ S_8 &= 0.779^{+0.030}_{-0.025} \\ \sigma_8 &= 0.810^{+0.061}_{-0.066}\,. \end{aligned} \tag{23}$$

Without performing baryon marginalization (yellow-green contours), the resulting marginal 2D posterior is still overlapping with the 1σ region of the case when introducing one parameter to account for baryonic uncertainty (blue), and with the result when the conservative scale cuts is adopted (yellow). By comparing the Ω_m–S_8 posterior distribution in the left-hand panel of Fig. 11 with the left-hand panel of Fig. 5, we find that the baryonic scenario as measured from DES is comparable to that of the Eagle mock-data simulation, for which we have learned that even without performing baryon marginalization when extending cosmic shear to 2.5′, the S_8 bias can still be within 0.5σ (see Section 3.1.1).

4.2 Informative prior on baryonic physics

The parameter constraints obtained when adopting our informative Q_1 prior is shown in the right-hand panel of Fig. 11 in the shaded grey contours in comparison with the baseline (blue contours) and the Y1 scale cut (yellow contours) results. The derived marginal 1D peak cosmological parameters are

$$\begin{aligned} \Omega_m &= 0.278^{+0.024}_{-0.034} \\ S_8 &= 0.788^{+0.018}_{-0.021} \\ \sigma_8 &= 0.821^{+0.052}_{-0.052}\,. \end{aligned} \tag{24}$$

Compared with the (averaged) error bars resulting from setting Y1 scale cuts, as shown in equation (22), we have around 11 per cent,

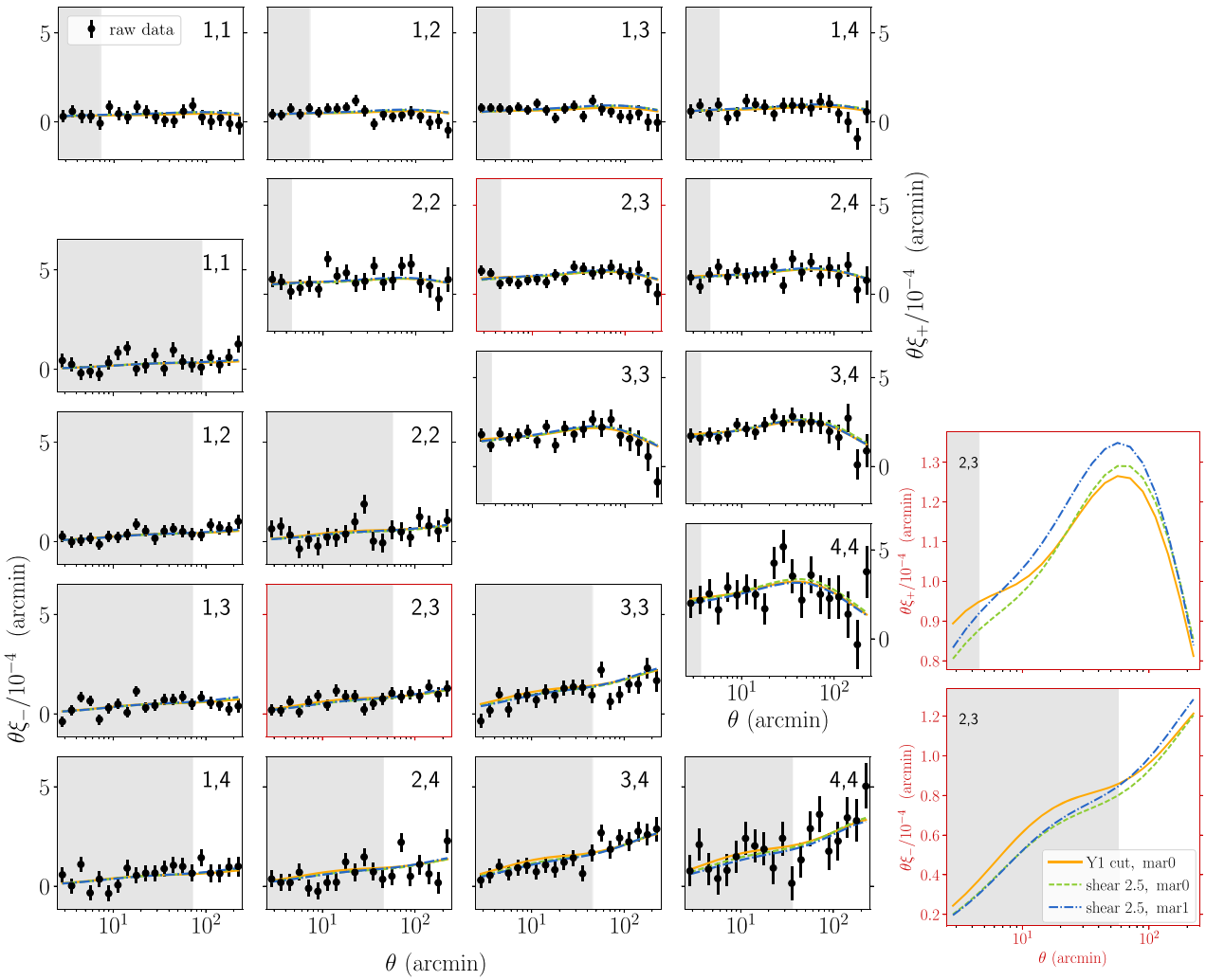

Figure 10. DES Y1 cosmic shear data vector (filled black points) and the best-fitting theoretical predictions from our baseline analyses. The yellow solid lines show the best-fitting model when the original Y1 scale cut is applied; the Y1 discarded data points are highlighted in the shaded grey regions. The blue dash-dotted (yellow-green dashed) lines indicate the best-fitting model when extending cosmic shear data points to 2.5 arcmin, and marginalizing over 1 (0) PC amplitude to account for baryons. For clarity, the two panels on the right show the (2, 3) tomographic bin with the rescaled vertical axis to better illustrate the differences between the models. The χ^2 information of these best-fitting models is summarized in Table 2.

Table 2. Goodness-of-fit for 3 × 2pt data for the best-fitting models (maximum likelihood point sampled in a chain), and the summary of constraints on the 1D peak value of Ω_m, S_8, and σ_8. The comparison between the cosmic shear data vectors and the model predictions for these best-fitting models are shown in Fig. 10. The first row lists the χ^2 values. The second and the third rows summarize the effective number of parameters (parameters subjected to wide priors), and the effective degrees of freedom for each of the analyses settings. The fourth row shows the reduced χ^2 values computed, with their errors provided in the fifth row. To understand whether a specific model is a good description of the data, in the sixth row, we derive the p-values based on the χ^2 distribution. A p-value > 0.05 indicates that the data are compatible with the model prediction within the error.

	Y1 cut	Shear to 2.5 arcmin (no baryon marginalization)	Shear to 2.5 arcmin (mar. Q_1)	DES (shear 2.5 arcmin, mar. Q_1) + Planck EE+BAO
Best-fitting χ^2	502	674	675	678
Effective $N_{\rm par}$	12	12	13	8
Effective d.o.f	443	618	617	622
Reduced χ^2	1.128	1.091	1.094	1.090
$\sqrt{2/\text{d.o.f}}$	0.067	0.057	0.057	0.057
p-value	0.027	0.059	0.053	0.059
Ω_m	$0.268^{+0.034}_{-0.031}$	$0.284^{+0.033}_{-0.027}$	$0.278^{+0.034}_{-0.031}$	$0.294^{+0.008}_{-0.006}$
S_8	$0.787^{+0.024}_{-0.025}$	$0.770^{+0.015}_{-0.017}$	$0.779^{+0.030}_{-0.025}$	$0.781^{+0.014}_{-0.015}$
σ_8	$0.831^{+0.060}_{-0.069}$	$0.784^{+0.047}_{-0.047}$	$0.810^{+0.061}_{-0.066}$	$0.788^{+0.010}_{-0.010}$

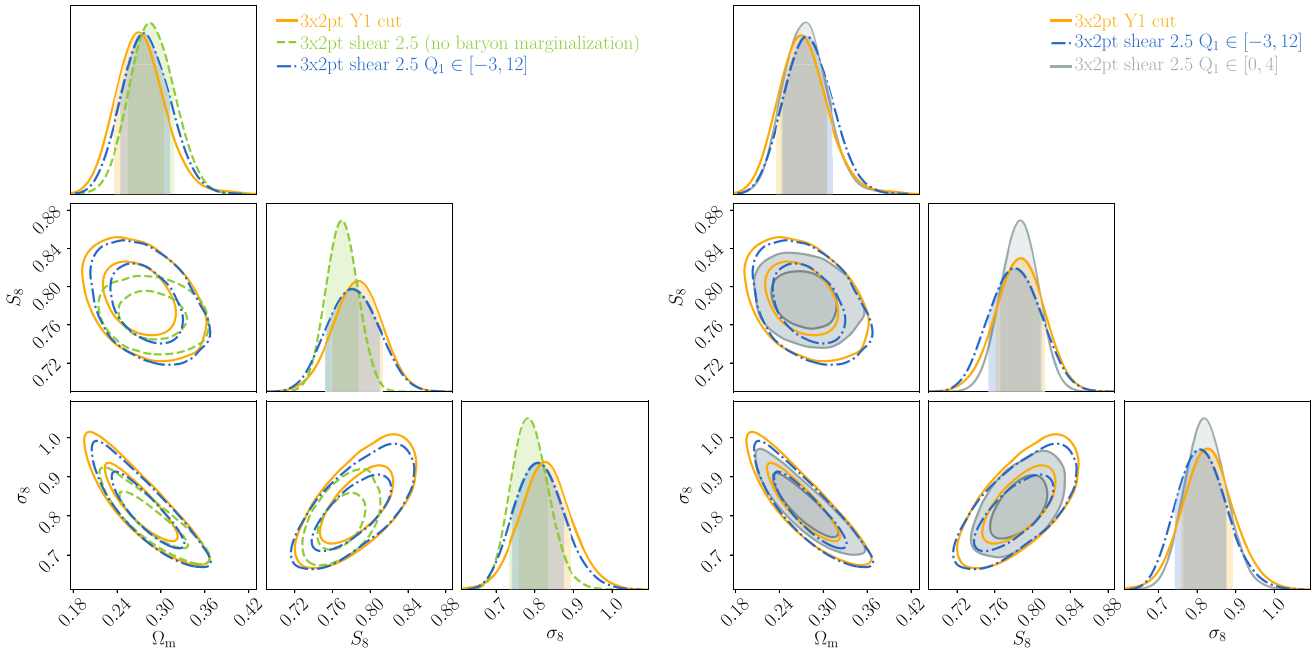

Figure 11. Cosmological parameter constraints on DES Y1 3 × 2pt data in ΛCDM. The yellow solid contours indicate the constraint when the original Y1 scale cut is applied. The blue dash-dotted contours show the result when adding extra small-scale cosmic shear data, and with uncertainties of baryon being marginalized with non-informative Q_1 prior. Left-hand panel: The yellow-green dashed contours indicate the result with small-scale cosmic shear data added but without performing baryon marginalization. Right-hand panel: The shaded grey contours indicate the case when extending cosmic shear to 2.5 arcmin while adopting informative prior on the first PC amplitude to perform baryon marginalization.

20 per cent, 19 per cent improvements on the 1D marginalized 1σ error bars of $\Omega_{\rm m}$, S_8, σ_8, respectively. The derived improvements are consistent with what we have learned from simulated likelihood analyses (Section 3.2) before unblinding.

4.3 Non-informative neutrino prior

Due to concerns about parameter projection effects as discussed in Section 2.4.1, we adopt an informative prior on the sum of the neutrino mass parameter based on the external information from Planck Collaboration VI (2018) and BAO measurements (Beutler et al. 2011; Ross et al. 2015; Alam et al. 2017). Here we explore the cosmology results for different choices of neutrino priors between our baseline ($\Omega_\nu h^2 \in$ Flat(5×10^{-4}, 1.3×10^{-3})) and the non-informative case as adopted in the original Y1 analysis ($\Omega_\nu h^2 \in$ Flat(5×10^{-4}, 0.1)).

Fig. 12 shows that the adoption of non-informative neutrino priors (purple contours) results in a slight shift ($\sim 0.3\sigma$ in S_8) in the posterior distribution compared to the case of an informative neutrino prior, which matches our previous observation using simulated likelihood analyses (see Section 2.4.1) before unblinding. The best-fitting cosmological parameters when including cosmic shear small-scale information and when marginalizing over uncertainties in baryonic physics with a non-informative prior on the Q_1 parameter are as follows:

$$\begin{aligned} \Omega_{\rm m} &= 0.286^{+0.037}_{-0.032} \\ S_8 &= 0.771^{+0.026}_{-0.028} \\ \sigma_8 &= 0.788^{+0.055}_{-0.069}\,. \end{aligned} \tag{25}$$

4.4 Constraints with external data

We compare our baseline DES measurements to external data from Planck (Planck Collaboration VI 2018) and BAO measurements

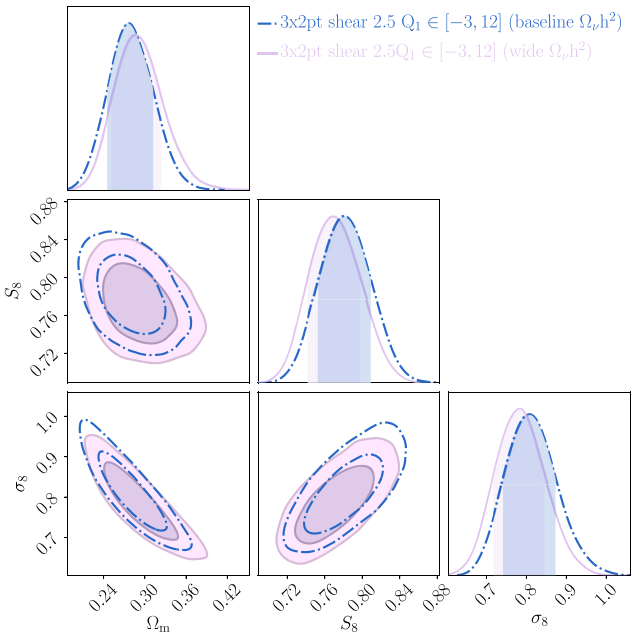

Figure 12. The effect of neutrino priors on the cosmological parameter constraints with DES Y1 3 × 2pt data in ΛCDM . The blue contours show our baseline result as in Fig. 11, for which an informative neutrino prior is applied ($\Omega_\nu h^2 \in$ Flat(5×10^{-4}, 1.3×10^{-3})). The shaded purple contours indicate the result when adopting a non-informative prior on neutrinos as the original Y1 analysis ($\Omega_\nu h^2 \in$ Flat(5×10^{-4}, 0.1)).

(Beutler et al. 2011; Ross et al. 2015; Alam et al. 2017). The main motivation is to use external information to tighten constraints on cosmology, and increase our constraining power on baryonic physics (see Section 5).

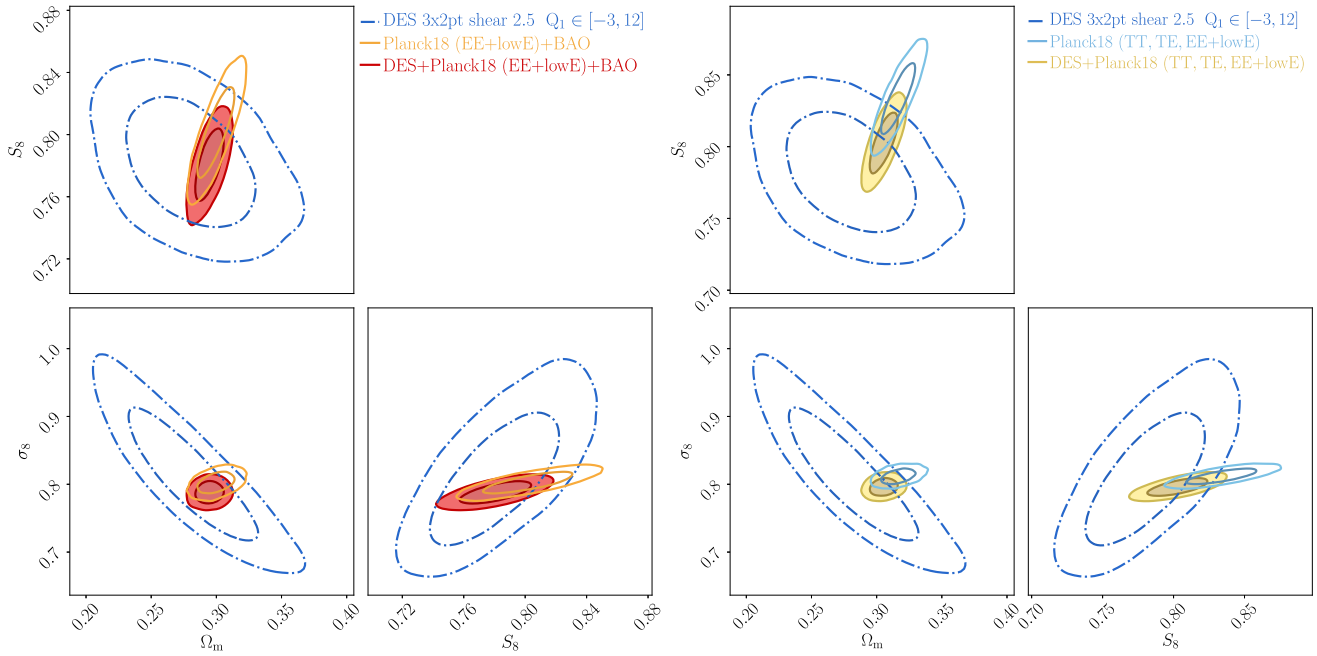

Figure 13. DES+Planck ΛCDM cosmology constraints. The blue contours show our baseline analysis on DES 3 × 2pt data with small-scale cosmic shear data included in the analysis, and marginalizing over the Q_1 parameter to account for baryon uncertainty. The orange contours in the left-hand panel show the Planck EE+lowE+BAO constraints, and the light blue contours in the right-hand panel display the Planck TT, TE, EE+lowE results from Planck Collaboration VI (2018). Within 2σ, the Planck contours are in agreement with our baseline DES result. The shaded contours present the joint constraints from our baseline DES analyses with the information from the Planck results.

We have considered two likelihood chains from the baseline DR18 Planck analyses[11]: the CMB polarization auto power spectra combined with BAO (referred to as EE+lowE+BAO), and the joint CMB temperature and polarization auto- and cross-power spectra (referred to as TT, TE, EE+lowE).

Our primary choice is the Planck EE+BAO likelihood, motivated by its high level of statistical consistency with DES Y1, as shown in the left-hand panel of Fig. 13. We compute 5D parameter covariance in Ω_m, A_s, n_s, Ω_b, and h from the Planck EE+BAO posterior distribution, and then rerun the DES Y1 data by adopting informative 5D Gaussian priors on these cosmological parameters. We have confirmed that the posterior of the Planck chains can be well-described with a multidimensional Gaussian fit out to the 4σ level. The shaded red contours in Fig. 13 present the combined result. The χ^2 analysis on the sampled maximum likelihood model indicates that our model prediction is consistent with the data (see Table 2).

The 1D marginal constraints are

$$\begin{aligned} \Omega_m &= 0.294^{+0.008}_{-0.006} \\ S_8 &= 0.781^{+0.014}_{-0.015} \\ \sigma_8 &= 0.788^{+0.010}_{-0.010}. \end{aligned} \tag{26}$$

There is ~ 50 per cent improvement in the S_8 constraint after adopting informative cosmological priors.

We also compare our DES Y1 analysis with the Planck TT, TT, TE constraint (light blue contours of Fig. 13). With the addition of Planck CMB temperature information, the CMB constraints reveal hints of tension with several ongoing weak lensing experiments (Hildebrandt et al. 2017; Abbott et al. 2018; Hikage et al. 2019), where the weak lensing results show lower values in the S_8 constraints. As shown in the right-hand panel of Fig. 13, although the two data sets are largely in agreement to within the 95 per cent confidence level in their 2D posterior constraints, the marginal 1D S_8 constraints differ by more than 1σ (see summary in Fig. 9).

[11] 2018 Planck Cosmological parameters and MCMC chains

5 BARYON CONSTRAINTS FROM DES Y1 DATA

In this section, we present the constraints on baryonic physics in terms of the first PC amplitude Q_1, which captures the most dominant features of baryonic effects on the clustering of the matter distribution. We first discuss the baryonic physics constraints from DES alone, and then increase the constraining power by combining DES with external data from Planck and BAO measurements.

5.1 DES-only information

The blue curves shown in Fig. 14 indicate the joint constraints of Q_1 and S_8 from the DES 3 × 2pt data with cosmic shear extended down to 2.5 arcmin, and with parameter priors listed in Table 1. As discussed in Appendix A, there is a significant positive correlation between Q_1 and S_8.

The marginal 1D Q_1 posterior for our baseline result is presented in the blue curve of Fig. 15. We find that Q_1 is constrained to be in the range:

$$\begin{aligned} -1.66 < Q_1 < 3.34 \quad &(68 \text{ per cent, DES Y1}) \\ -2.96 < Q_1 < 5.63 \quad &(95 \text{ per cent, DES Y1}) . \end{aligned} \tag{27}$$

We can rule out the cosmo-OWLS scenario with AGN minimum heating temperature setting at $10^{8.7}$ K at $\sim 2.1\sigma$ with DES alone. This conclusion still holds (and the Q_1 posterior remains similar) for the analysis when adopting the original Y1 wide neutrino prior, as shown in the purple contours of Fig. 14.

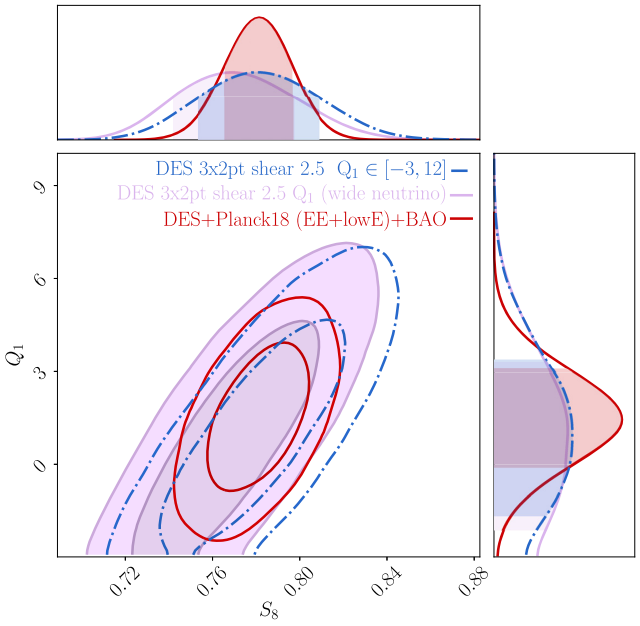

Figure 14. Joint constraints on cosmological parameters and baryonic physics. The blue contours show the Q_1–S_8 constraints from the baseline analyses of DES 3 × 2pt data with cosmic shear measured down to 2.5 arcmin and with the baseline priors detailed in Table 1. The analysis setting for the purple contour is the same as the blue contour, except for the adoption of the non-informative neutrino prior as used in the original Y1 analysis. The red contours present the results when adopting informative cosmology priors from the external information of Planck 2018 EE+BAO.

As concluded in Section 3.3 from the results of the simulated likelihood analyses, we expect a $\lesssim 0.5\sigma$ shift in the peak of the marginal 1D Q_1 posterior distribution due to the parameter projection effects driven by the degeneracies between parameters.

In the next section, we explore the constraints on baryonic physics when including cosmological information from external data sets.

5.2 Adopting cosmological parameter priors from external data sets

As discussed in Section 4.4, we adopt the Planck EE+BAO likelihood as the primary source of prior information on cosmological parameters, due to its consistency with DES Y1 data.

With the inclusion of the Planck EE+BAO information, the constraints on cosmology improve by ~ 38 per cent in the 68 per cent confidence interval of Q_1 (see the red curve in Fig. 15), which is further quantified as

$$\begin{aligned} -0.06 < Q_1 < 3.04 \quad &(68 \text{ per cent}, \quad \text{DES+Planck EE+BAO}) \\ -1.68 < Q_1 < 4.66 \quad &(95 \text{ per cent}, \quad \text{DES+Planck EE+BAO}) \,. \end{aligned} \tag{28}$$

With the tighter constraining power, the cosmo-OWLS $10^{8.7}$ K scenario is disfavoured by $\sim 2.8\sigma$.

We note that when claiming the cosmo-OWLS ($10^{8.7}$ K) scenario is ruled out at $>2\sigma$ significance with DES Y1 data, this does not mean that we are ruling out setting the subgrid physical parameter of $\Delta T_{\rm heat}$ at $10^{8.7}$ K when running the hydrodynamical simulations at $>2\sigma$. Rather, we are simply ruling out the level of baryonic suppression in the cosmic shear signal exhibited by that simulation, as shown in Fig. 16 based on the empirical PCA framework. As discussed in Section 2.1.3, for baryon models with realistic physical parametrizations, such as the baryon correction model of Schneider et al. (2020a) or subgrid physical parameters in hydrodynamical simulations, the baryonic physics parameters controlling halo density profiles couple to the cosmic baryon fraction $f_{\rm b} = \Omega_{\rm b}/\Omega_{\rm m}$. As revealed in the trend of power spectrum ratios in fig. 6 of van Daalen et al. (2020), at fixed $\Delta T_{\rm heat}$ value in hydrosimulations, the baryonic suppression features also vary slightly at different cosmologies with different $f_{\rm b}$. For a Universe with lower $f_{\rm b}$, higher AGN heating temperature is needed to generate the same amount of suppression feature compared with a Universe with higher $f_{\rm b}$.

Finally, we link back to the physical effect of Q_1, which is best demonstrated by looking at the suppression of the amplitude of cosmic shear correlation functions (Fig. 16, also see Fig. A2). In Fig. 16 we convert the 1σ constraints on Q_1 to cosmic shear model vectors via equation (19), and present the ratio of the baryonic physics-included model with respect to the DMO-based theoretical model. We use the pair of tomographic bins (2, 3) to demonstrate the effects of baryons. The thick lines indicate the results when setting the Q_1 value at the 1D marginal peak the posteriors, as indicated in the text in the right-hand panel of Fig. 14. The other baryonic scenarios are also overplotted in thinner curves for comparison. The figure shows the effect of the baryonic effects on the shear–shear observables, and can be compared to Fig. 1 where we show the effects on the matter power spectrum. The shear–shear correlation function measured on a range of scales and tomographic bins can constrain both the spatial and temporal evolution of the baryonic effects.

6 DISCUSSION AND SUMMARY

Small-scale information in galaxy imaging surveys has substantial statistical power to improve cosmological constraints. But conventional cosmological analyses discard this information to avoid biased inference of cosmology due to an insufficient theoretical description of the sources of astrophysical and observational systematic uncertainty on these scales.

The effects of baryonic physics constitute the dominant source of uncertainty at small scales in the matter power spectrum (van Daalen et al. 2011, 2020). A variety of modelling and mitigation strategies have been proposed in the literature to account for the complicated mechanisms involved, such as baryonic feedback and cooling processes (see Chisari et al. 2019 for a review of existing baryon mitigation methods).

To enable robust inference of cosmological parameters, it is desirable to find a minimal parametrization that accurately captures the effects of baryonic physics on the observables and to have stringent priors on these parameters. Principal component analysis (PCA) of the cosmological observables, derived from a set of hydrodynamical simulations that span the range of allowed baryonic scenarios, is one of the most promising avenues to obtain such a minimal parametrization (Eifler et al. 2015; Kitching et al. 2016; Mohammed & Gnedin 2018; Huang et al. 2019).

In this paper, we employ the hydrodynamical simulation-based PCA method to parametrize baryonic effects in the Dark Energy Survey Year 1 data. We find that one principal component is sufficient to capture the range of baryonic physics at the level of DES Y1 statistical constraining power. We include the amplitude of this PC, Q_1, as an additional parameter in our likelihood analysis. The magnitude of Q_1 reflects the strength of baryonic feedback, with larger Q_1 values corresponding to a stronger suppression of small-scale cosmic shear correlation functions (see Fig. A2).

Previous DES multiprobe analyses (e.g. Abbott et al. 2018, 2019) impose stringent scale cuts to ensure that the analysis is unaffected by

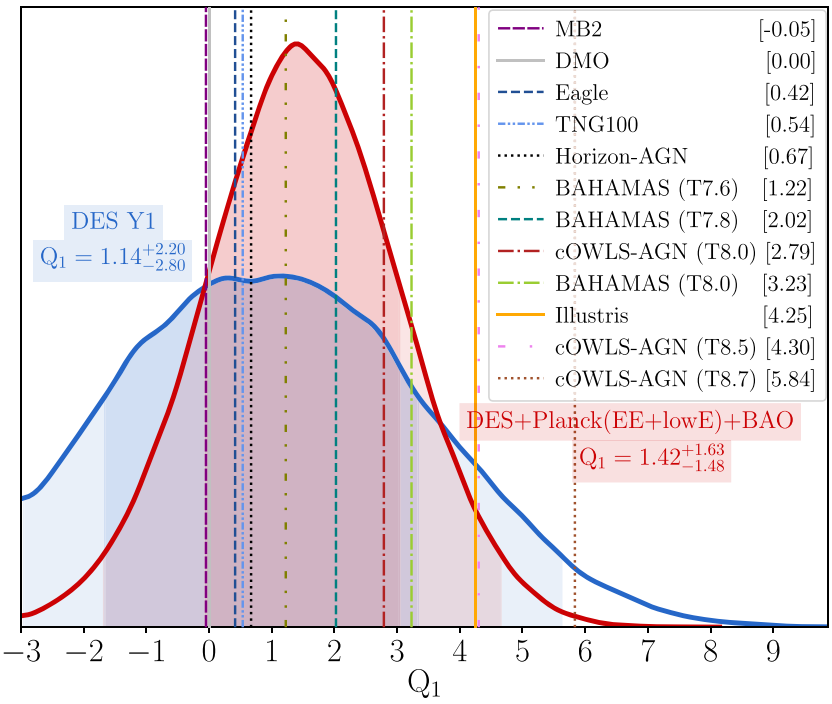

Figure 15. Constraints on the baryonic parameter Q_1 from the DES Y1 information only (blue) and with the adoption of the Planck EE+BAO cosmological prior (red). The shaded bands highlight the 68 per cent and the 95 per cent confidence intervals of the constraints. The representative Q_1 values from various baryonic scenarios are shown in the legend and are overplotted as vertical lines. The most extreme baryonic scenario, the cosmoOWLS-AGN with minimum AGN heating temperature of $10^{8.7}$K, is excluded at above 2σ level by the real data.

baryonic physics; this had the largest effect on the range of scales used by cosmic shear. The inclusion of baryonic effects in our theoretical model for the observables allows us to relax these scale cuts and to include scales as small as 2.5 arcmin from cosmic shear in the analysis. We otherwise follow the DES Year 1 systematics modelling and mitigation strategy, except for adding an informative neutrino mass prior based on findings by the Planck satellite mission. The reduced range in varying neutrino mass avoids parameter volume effects in the DES analysis (see Section 2.4.1).

Our joint analysis of baryonic physics and cosmology (in combination with the other DES systematics parameters) yields $S_8 = 0.779^{+0.030}_{-0.025}$ if we allow for self-calibration of Q_1. When we restrict the range of Q_1 such that AGN feedback stronger than the level of Illustris (Genel et al. 2014; Vogelsberger et al. 2014) is excluded, we get $S_8 = 0.788^{+0.018}_{-0.021}$ (see the right-hand panel of Fig. 11).

We proceed to combine DES Y1 with data from the latest Planck mission analysis (Planck Collaboration VI 2018). However, we exclude the Planck temperature power spectrum information due to an abundance of caution as to whether these data sets might be in tension (Adhikari & Huterer 2019; Garcia-Quintero et al. 2019; Park & Rozo 2019). We instead use the Planck EE+lowE+BAO chain as described in Planck Collaboration VI (2018), which also includes BAO measurements from the BOSS DR12, 6DF and MGS survey. Our joint DES Y1+Planck+BAO analysis yields $S_8 = 0.781^{+0.014}_{-0.015}$ (see Fig. 13).

We emphasize that the main goal of this paper is not to find the tightest possible constraints on cosmological parameters, but rather unbiased constraints on baryonic physics with cosmological parameters being allowed to vary. We find the baryon parameter $Q_1 = 1.14^{+2.20}_{-2.80}$ for DES Y1 only and $Q_1 = 1.42^{+1.63}_{-1.48}$ for DES+Planck EE+BAO (see Fig. 15), which allows us to exclude one of the most extreme AGN feedback hydrodynamical scenario, cosmo-OWLS AGN ($T = 10^{8.7}$K), at $\sim 2.8\sigma$.

Among the 11 hydrodynamical scenarios in our pool, the default BAHAMAS simulation (minimum AGN heating temperature at $T = 10^{7.8}$ K) is perhaps the best-calibrated baryon scenario. Not only is it tuned to reproduce the galaxy stellar mass function, but it also has adjusted feedback parameters so that the halo hot gas mass fractions match those from the observations (McCarthy et al. 2017). The 1σ region of our Q_1 posterior constraint likewise includes the default BAHAMAS scenario.

Constraining the strength of baryon feedback is also important to understand whether it can serve as a possible explanation for the 'lensing-is-low' effect, i.e. the fact that the observed galaxy–galaxy lensing signal is low by ∼20–40 per cent compared to predictions from N-body+HOD mocks, at fixed clustering signal (Leauthaud et al. 2017). As discussed in Lange et al. (2019), the IllustrisTNG scenario can account for ∼10 per cent of the suppression signal and the stronger feedback scenario of Illustris can reach to ∼15 per cent (cf. Fig. 15 for our constraints on these scenarios). While our constraints currently lack the constraining power to make definite statements on ruling out baryonic effects as a potential explanation for the lensing-is-low signal, we expect that future analyses, e.g. using DES Y3 data, will be very interesting in that regard.

Generally speaking, our resulting Q_1 posterior distribution indicates a preference for moderate to weak baryonic feedback, which is consistent with previous cosmic shear constraints of Joudaki et al. (2017) for an analysis on CFHTLenS data, and with MacCrann et al. (2017) for the DES SV data (Abbott et al. 2016) using `HMCODE` (Mead et al. 2015). Recent constraints from the cosmic shear

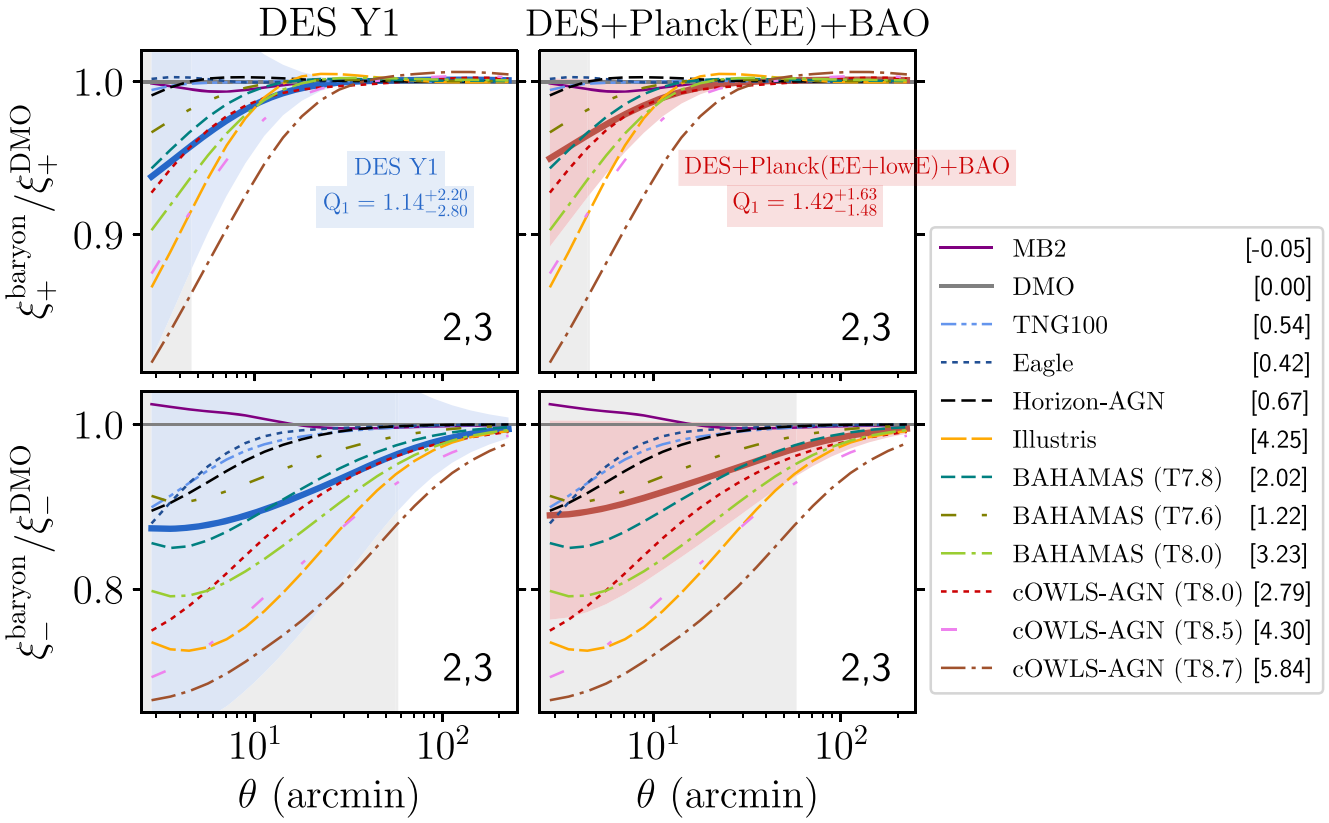

Figure 16. Quantifying the strength of baryonic feedback using the Q_1 parameter. Here we show the ratio of cosmic shear correlation functions for the 11 baryonic scenarios with respect to the theoretical (DMO) predictions (thinner curves), using the pair of tomographic bins (2,3) as an demonstration. The grey bands highlight the angular scales that were excluded in the fiducial DES Y1 analysis, but are now included in this work. The thick curves depict the baryonic features using the best-fitting Q_1 constraints from our baseline DES Y1 analysis (left-hand panel, blue) and the result with the adoption of cosmological priors from Planck 2018 EE+BAO information (right-hand panel, red). The shaded bands highlight the 1σ region of our data constraints. The numbers in the coloured-shaded legend are the best-fitting Q_1 values from data constraints, and the representative Q_1 values from various baryonic scenarios are also provided in the right-hand-side legend.

KiDS-VIKING 450^2 degree field (Hildebrandt et al. 2020) analysed by Yoon & Jee (2020) also derive a weak signal of baryonic feedback based on the baryonic model of HMCODE. They found that the baryon suppression signal is consistent with the DMO scenario within 1.2σ significance with the KiDS data alone, and is at 2.2σ level deviation from DMO under the assumption of the WMAP9 cosmology. On the contrary, based on the analysis of the DLS (Deep Lens Survey; Jee et al. 2016) Fourier space galaxy-mass and galaxy–galaxy power spectra, Yoon et al. (2019) report a preference for strong baryonic feedback that is more extreme than that predicted by OWLS-AGN.[12] The difference in the resulting baryon constraints could be the result of Yoon et al. (2019) adopting a linear galaxy bias model, which may not be a sufficient assumption to interpret the data points to scales as small as $\ell \sim 2000$ (see fig. 7 of Krause et al. 2017 for the determination of scale cuts in DES Y1 galaxy lensing and galaxy clustering observables to avoid the impact of non-linear galaxy bias).

Although baryonic effects are the dominant systematic uncertainty on small-scale cosmic shear measurements, there are other systematics that will likely become important in future, more constraining analyses. Contributions from third-order corrections of the shear two-point correlations, such as reduced shear (Shapiro 2009) and magnification bias (Schmidt et al. 2009) effects are estimated to produce a ~ 2 per cent fractional difference in the observables of ξ_+ and ~ 5 per cent in ξ_- at 2.5 arcmin. If not accounted for, they would lead to a $\sim 1\sigma$-level bias in the constraint of the HMCODE baryon parameter (Mead et al. 2015) under a DES Y5-like data quality, according to MacCrann et al. (2017) (see their figs 5 and 7). The choice of IA models can also affect baryon constraints in future more constraining data sets, given that IA and cosmological parameters are degenerate (see e.g. fig. 4 of MacCrann et al. 2017). For Y1, switching from the simple NLA model to the full tidal alignment and tidal torquing (TATT) model (Blazek et al. 2019) leads to a slight shift of $\sim 0.5\sigma$ in S_8, but overall the resulting likelihoods are still in agreement within Y1 errors (Samuroff et al. 2018; Troxel et al. 2018a). The improved data quality in forthcoming data sets will likely mean that discrepancies induced from these small-scale systematics will become non-negligible, and will require extra efforts to extract precise joint constrains on both cosmology and baryonic physics.

The ongoing KiDS and HSC analyses provide an excellent data set to get additional insights into discriminating between different baryonic physics scenarios. Moreover, future data sets from DES Year 3 and Year 6 will provide improved joint constraints on baryonic physics and cosmological parameters.

In the regime where effects of baryonic physics are causing the suppression of clustering power ($k \lesssim$ few $10\,{\rm Mpc}^{-1}h$), the properties of halo gas contain a wealth of information on baryon feedback

[12]The OWLS-AGN scenario is equivalent to the cosmoOWLS AGN scenario with T=$10^{8.0}$ K as indicated in the red line of Fig. 15

mechanisms. Observational probes such as X-ray, thermal, and kinetic Sunyaev–Zel'dovich measurements are directly sensitive to the distribution and the characteristics of gas content (e.g. Battaglia et al. 2017). Ultimately, utilizing information from both gas-sensitive observables and weak lensing provides the most promising avenue to constrain baryon feedback (Hojjati et al. 2017; Aricò et al. 2020; Debackere et al. 2020; Mead et al. 2020; Osato et al. 2020; Pandey, Baxter & Hill 2020; Schneider et al. 2020b)

As we prepare for future analyses of the Rubin Observatory LSST, *Euclid*, SPHEREx, and *Roman Space Telescope*, the information regarding which baryonic scenarios are already excluded by Stage III data is invaluable for the design of cosmology analysis pipelines, and simulation efforts in order to focus the computational power where it is needed most and in order to optimally analyse these future data sets.

ACKNOWLEDGEMENTS

We thank the internal reviewers from the DES collaboration for providing insightful comments and giving valuable feedback. HH and TE are supported by NASA ROSES ADAP, grant 16-ADAP16-0116 and Department of Energy Cosmic Frontier program, grant DE-SC0020215. RM is supported by the Department of Energy Cosmic Frontier program, grant DE-SC0010118.

DATA AVAILABILITY STATEMENT

The observational data underlying this article are publicly available in the official DES Year 1 key project website at https://des.ncsa.illinois.edu/releases/y1a1/key-products.

REFERENCES

Abbott T. et al., 2016, Phys. Rev. D, 94, 022001
Abbott T. M. C. et al., 2018, Phys. Rev. D, 98, 043526
Abbott T. M. C. et al., 2019, Phys. Rev. D, 100, 023541
Adhikari S., Huterer D., 2019, J. Cosmol. Astropart. Phys., 2019, 036
Alam S. et al., 2017, MNRAS, 470, 2617
Aricò G., Angulo R. E., Hernández-Monteagudo C., Contreras S., Zennaro M., Pellejero-Ibañez M., Rosas-Guevara Y., 2020, MNRAS, 495, 4800
Asgari M. et al., 2020, A&A, 634, A127
Barreira A., Nelson D., Pillepich A., Springel V., Schmidt F., Pakmor R., Hernquist L., Vogelsberger M., 2019, MNRAS, 488, 2079
Battaglia N., Ferraro S., Schaan E., Spergel D. N., 2017, J. Cosmol. Astropart. Phys., 2017, 040
Beutler F. et al., 2011, MNRAS, 416, 3017
Blas D., Lesgourgues J., Tram T., 2011, J. Cosmol. Astropart. Phys., 2011, 034
Blazek J. A., MacCrann N., Troxel M. A., Fang X., 2019, Phys. Rev. D, 100, 103506
Brown M. L., Taylor A. N., Hambly N. C., Dye S., 2002, MNRAS, 333, 501
Chisari N. E. et al., 2018, MNRAS, 480, 3962
Chisari N. E. et al., 2019, Open J. Astrophys., 2, 4
Copeland D., Taylor A., Hall A., 2018, MNRAS, 480, 2247
Dai B., Feng Y., Seljak U., 2018, J. Cosmol. Astropart. Phys., 2018, 009
Davis C. et al., 2017, preprint (arXiv:1710.02517)
Debackere S. N. B., Schaye J., Hoekstra H., 2020, MNRAS, 492, 2285
DeRose J. et al., 2019, ApJ, 875, 69
Doré O. et al., 2014, preprint (arXiv:1412.4872)
Dubois Y. et al., 2014, MNRAS, 444, 1453
Eifler T., 2011, MNRAS, 418, 536
Eifler T., Krause E., Dodelson S., Zentner A. R., Hearin A. P., Gnedin N. Y., 2015, MNRAS, 454, 2451
Eifler T. et al., 2020a, preprint (arXiv:2004.04702)
Eifler T. et al., 2020b, preprint (arXiv:2004.05271)
Elvin-Poole J. et al., 2018, Phys. Rev. D, 98, 042006
Fang X., Krause E., Eifler T., MacCrann N., 2020, J. Cosmol. Astropart. Phys., 2020, 010
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Garcia-Quintero C., Ishak M., Fox L., Lin W., 2019, Phys. Rev. D, 100, 123538
Gatti M. et al., 2018, MNRAS, 477, 1664
Genel S. et al., 2014, MNRAS, 445, 175
Giblin B. et al., 2018, MNRAS, 480, 5529
Goodman J. et al., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Haider M., Steinhauser D., Vogelsberger M., Genel S., Springel V., Torrey P., Hernquist L., 2016, MNRAS, 457, 3024
Hamana T. et al., 2020, PASJ, 72, 16
Harnois-Déraps J., van Waerbeke L., Viola M., Heymans C., 2015, MNRAS, 450, 1212
Hearin A. P., Zentner A. R., Ma Z., 2012, J. Cosmol. Astropart. Phys., 2012, 034
Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, ApJ, 715, 104
Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 111
Heymans C. et al., 2006, MNRAS, 368, 1323
Hikage C. et al., 2019, PASJ, 71, 43
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Hildebrandt H. et al., 2020, A&A, 633, A69
Hinshaw G. et al., 2013, ApJS, 208, 19
Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526
Hojjati A. et al., 2017, MNRAS, 471, 1565
Hoyle B. et al., 2018, MNRAS, 478, 592
Huang H.-J., Eifler T., Mandelbaum R., Dodelson S., 2019, MNRAS, 488, 1652
Huff E., Mandelbaum R., 2017, preprint (arXiv:1702.02600)
Huterer D., Takada M., 2005, Astropart. Phys., 23, 369
Huterer D., Takada M., Bernstein G., Jain B., 2006, MNRAS, 366, 101
Ivezić Ž. et al., 2019, ApJ, 873, 111
Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338
Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt S., Wittman D., 2016, ApJ, 824, 77
Joudaki S. et al., 2017, MNRAS, 465, 2033
Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349
Kitching T. D., Verde L., Heavens A. F., Jimenez R., 2016, MNRAS, 459, 971
Krause E. et al., 2017, preprint (arXiv:1706.09359)
Krause E., Eifler T., 2017, MNRAS, 470, 2100
Kuijken K. et al., 2019, A&A, 625, A2
Lange J. U., Yang X., Guo H., Luo W., van den Bosch F. C., 2019, MNRAS, 488, 5771
Laureijs R. et al., 2011, Euclid Definition Study Report. preprint (arXiv:1110.3193)
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Leauthaud A. et al., 2017, MNRAS, 467, 3024
Lin C.-H., Harnois-Déraps J., Pospisil T., Mandelbaum R., Lee A. B., Singh S., 2020, MNRAS, 499, 2977
MacCrann N. et al., 2017, MNRAS, 465, 2567
Mandelbaum R. et al., 2018, PASJ, 70, S25
Marinacci F. et al., 2018, MNRAS, 480, 5113
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958
Mead A. J., Heymans C., Lombriser L., Peacock J. A., Steele O. I., Winther H. A., 2016, MNRAS, 459, 1468
Mead A. J., Tröster T., Heymans C., Van Waerbeke L., McCarthy I. G., 2020, A&A, 641, A130
Mohammed I., Gnedin N. Y., 2018, ApJ, 863, 173

Mohammed I., Martizzi D., Teyssier R., Amara A., 2014, preprint (arXiv:1410.6826)
Mummery B. O., McCarthy I. G., Bird S., Schaye J., 2017, MNRAS, 471, 227
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, Comput. Astrophys. Cosmol., 6, 2
Osato K., Shirasaki M., Miyatake H., Nagai D., Yoshida N., Oguri M., Takahashi R., 2020, MNRAS, 492, 4780
Pandey S., Baxter E. J., Hill J. C., 2020, Phys. Rev. D, 101, 043525
Park Y., Rozo E., 2020, MNRAS, 499, 4638
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Pillepich A. et al., 2018, MNRAS, 475, 648
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck Collaboration XIII, 2016, A&A, 594, A13
Planck Collaboration VI, 2020, A&A, 641, A6
Prat J. et al., 2018, Phys. Rev. D, 98, 042005
Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, MNRAS, 449, 835
Rozo E. et al., 2016, MNRAS, 461, 1431
Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19
Samuroff S. et al., 2018, MNRAS, 475, 4524
Schaye J. et al., 2010, MNRAS, 402, 1536
Schaye J. et al., 2015, MNRAS, 446, 521
Schmidt F., Rozo E., Dodelson S., Hui L., Sheldon E., 2009, ApJ, 702, 593
Schneider A. et al., 2020b, J. Cosmol. Astropart. Phys., 2020, 020
Schneider A., Teyssier R., 2015, J. Cosmol. Astropart. Phys., 2015, 049
Schneider P., Eifler T., Krause E., 2010, A&A, 520, A116
Schneider A., Teyssier R., Stadel J., Chisari N. E., Le Brun A. M. C., Amara A., Refregier A., 2019, J. Cosmol. Astropart. Phys., 2019, 020
Schneider A., Stoira N., Refregier A., Weiss A. J., Knabenhans M., Stadel J., Teyssier R., 2020a, J. Cosmol. Astropart. Phys., 2020, 019
Seljak U., 2000, MNRAS, 318, 203
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148
Shapiro C., 2009, ApJ, 696, 775
Sheldon E. S., Huff E. M., 2017, ApJ, 841, 24
Simpson F., James J. B., Heavens A. F., Heymans C., 2011, Phys. Rev. Lett., 107, 271301
Simpson F., Heavens A. F., Heymans C., 2013, Phys. Rev. D, 88, 083510
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Springel V. et al., 2018, MNRAS, 475, 676
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Taylor P. L., Bernardeau F., Kitching T. D., 2018, Phys. Rev. D, 98, 083514
Taylor P. L., Bernardeau F., Huff E., 2021, Phys. Rev. D, 103, 043531
Tenneti A., Mandelbaum R., Di Matteo T., Kiessling A., Khandai N., 2015, MNRAS, 453, 469
Troxel M. A. et al., 2018a, Phys. Rev. D, 98, 043528
Troxel M. A. et al., 2018b, MNRAS, 479, 4998
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
van Daalen M. P., McCarthy I. G., Schaye J., 2020, MNRAS, 491, 2424
van Uitert E. et al., 2018, MNRAS, 476, 4662
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Yoon M., Jee M. J., Tyson J. A., Schmidt S., Wittman D., Choi A., 2019, ApJ, 870, 111
Yoon M., Lee W., Jee M. J., Finner K., Smith R., Kim J.-W., 2020, ApJ, 903, 151
Zentner A. R., Rudd D. H., Hu W., 2008, Phys. Rev. D, 77, 043507
Zentner A. R., Semboloni E., Dodelson S., Eifler T., Krause E., Hearin A. P., 2013, Phys. Rev. D, 87, 043509
Zuntz J. et al., 2015, Astron. Comput., 12, 45
Zuntz J. et al., 2018, MNRAS, 481, 1149

APPENDIX A: THE INTERPLAY BETWEEN BARYONS, COSMOLOGY, AND OTHER SYSTEMATIC PARAMETERS

In this appendix, we investigate the degeneracies of the baryonic physics parameter Q_1 with cosmological and other nuisance parameters (see Table 1).

To quantify the level of parameter degeneracies, we compute correlation coefficients of the marginalized 2D posterior distributions between Q_1 with all the other parameters. The parameter correlation $\mathbf{Corr}^{ij}_{\rm par}$ is computed via:

$$\mathbf{Corr}^{\rm ij}_{\rm par} = \mathbf{C}^{ij}_{\rm par}/\sqrt{\mathbf{C}^{ii}_{\rm par}\mathbf{C}^{jj}_{\rm par}}\,, \tag{A1}$$

with the parameter covariance matrix computed as

$$\mathbf{C}^{ij}_{\rm par} = \frac{1}{N-1}\sum_{k=1}^{N}(\boldsymbol{\theta}^{ik} - \langle\boldsymbol{\theta}^{i}\rangle)(\boldsymbol{\theta}^{jk} - \langle\boldsymbol{\theta}^{j}\rangle)\,. \tag{A2}$$

The $\langle\boldsymbol{\theta}^{i}\rangle$ indicates the mean of the i-th parameter, and $k \in [1, N]$ is the index running over the first 90 per cent higher likelihood steps in the MCMC chain. We discard the 10 per cent of the MCMC samples with the lowest likelihood values when deriving the parameter covariance, in order to decrease the effects from samples distributed far away from the high likelihood region.

Using the likelihood simulation chain with the DMO scenario as mock data, in Fig. A1 we display the posterior distributions between Q_1 and parameters that are significantly correlated with it.

To understand the trends of parameter degeneracies from the MCMC, in Fig. A2, we plot the fractional changes in model data vectors ($\frac{\boldsymbol{M}-\boldsymbol{M}_{\rm fid}}{\boldsymbol{M}_{\rm fid}}$) when varying individual parameters to 1σ above (solid lines) or below (dash lines) from their fiducial values listed in Table 1. The positive correlation trend between Q_1 and $n_{\rm s}$ is clearly due to their opposing effects on the model vector, especially in ξ_-, as shown in the second column of Fig. A2 (red versus purple curves).

The significant positive correlation between Q_1 and S_8 explains the tendency towards parameter projection effects discussed in Section 3.1 and Section 3.3. This correlation is straightforward to understand. An increase in S_8 boosts the overall amplitude of matter clustering, whereas increasing the amount of feedback suppresses the clustering signal on small scales. The opposite correlation directions between Q_1 with $\Omega_{\rm m}$ (-0.33) and with σ_8 (0.54) are driven by the significant negative coupling between $\Omega_{\rm m}$ and σ_8.

Regarding the negative correlation observed between Q_1 and the galaxy bias parameters,[13] this correlation is actually driven by the common degeneracies of Q_1 and $b^i_{\rm g}$ with S_8. Ideally there should be almost no correlation between Q_1 and $b^i_{\rm g}$ because the variation of Q_1 is mostly affecting the cosmic shear observables, whereas the galaxy bias parameters only affect galaxy–galaxy lensing and galaxy clustering (see Fig. A2). However, when cosmology is allowed to vary, Q_1 and $b^i_{\rm g}$ appear correlated because of their correlation with cosmological parameters (mostly in $\Omega_{\rm m}$ and σ_8).[14]

[13] In Fig. A1 we pick the galaxy bias parameter of the second tomographic bin $b^2_{\rm g}$ as a demonstration; however, the parameter correlations between Q_1 and other tomographic bins $b^i_{\rm g}$ show similar results.

[14] The same logic also explains why there are tight correlations between pairs of bias parameters ($b^i_{\rm g}$ and $b^j_{\rm g}$). In principle, each bias parameter governs different portions of the model vector via the tomography bin division (if there are no photo-z uncertainties when dividing lens galaxies), and thus should have no correlation. The observed high correlations among the bias parameters are driven by the fact that they have a similar impact on $\Omega_{\rm m}$ and σ_8.

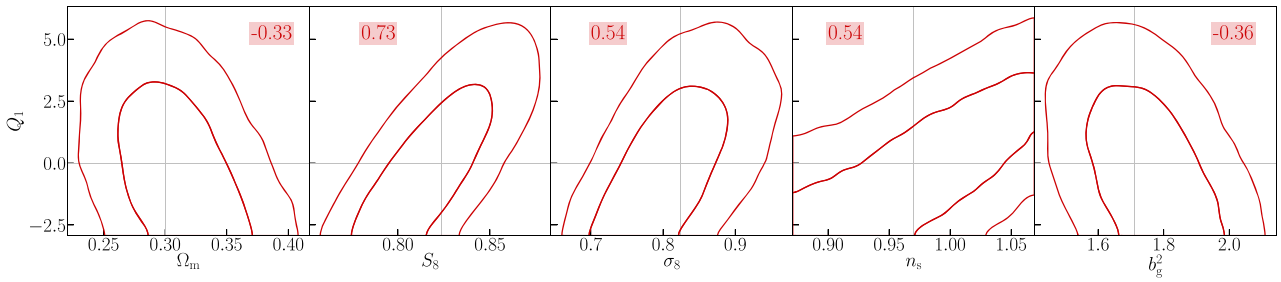

Figure A1. Marginalized 2D posterior distributions between Q_1 with parameters that show apparent parameter degeneracies, from the likelihood simulated analysis with the DMO mock data vector as input. The fiducial parameter values are indicated by the cross grey lines. The correlation coefficients are provided in the legend.

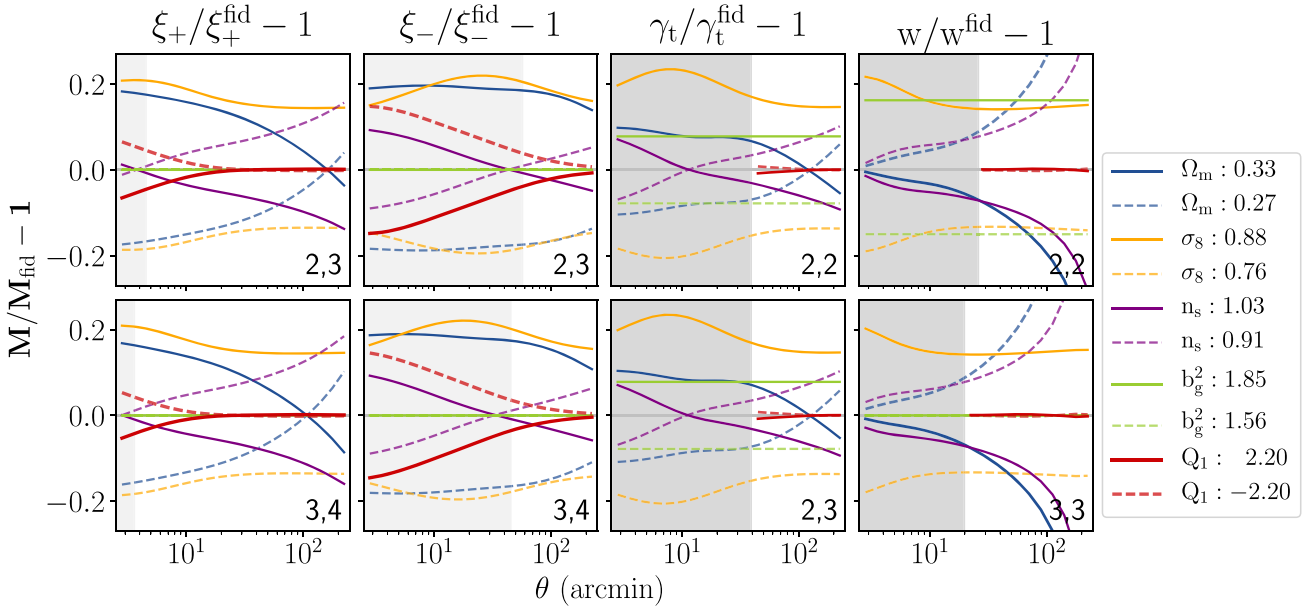

Figure A2. The fractional changes in model vectors when varying each of the individual cosmological or baryonic parameters to 1σ above (solid lines) or below (dash lines) their fiducial values listed in Table 1. We only select two tomographic bins for each of the observables as demonstration, with the bin information indicated on the bottom right corner of each panel. For galaxy–galaxy lensing, the first number is for lens tomographic bin; the second number for source. The darker grey bands in galaxy–galaxy lensing and clustering panels mark data points that are excluded throughout this work. The lighter grey bands in the cosmic shear panels highlight data points that are excluded in the original DES Y1 analysis, but are now included in this work. When varying cosmological parameters, we carefully adjust the A_s parameter in COSMOLIKE in order to keep σ_8 fixed.

[1]*Steward Observatory/Department of Astronomy, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[2]*McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[3]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[4]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[5]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[6]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain*
[7]*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*
[8]*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*
[9]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[10]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[11]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[12]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[13]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona) Spain*
[14]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[15]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain*
[16]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[17]*Kavli Institute for Particle Astrophysics and Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA*
[18]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[19]*Département de Physique Théorique and Center for Astroparticle Physics, Université de Genève, 24 quai Ernest Ansermet, CH-1211 Geneva, Switzerland*
[20]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

[21]*Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland*
[22]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[23]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, D-81679 München, Germany*
[24]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*
[25]*Department of Physics, Duke University Durham, NC 27708, USA*
[26]*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*
[27]*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
[28]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP 05314-970, Brazil*
[29]*Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil*
[30]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*
[31]*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA*
[32]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[33]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[34]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[35]*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[36]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*
[37]*National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[38]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil*
[39]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid 04200, Spain*
[40]*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany*
[41]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[42]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[43]*Center for Astrophysics, Harvard and Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[44]*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
[45]*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
[46]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*
[47]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[48]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[49]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton BN1 9QH, UK*
[50]*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[51]*Brandeis University, Physics Department, 415 South Street, Waltham, MA 02453, USA*
[52]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

This paper has been typeset from a TEX/LATEX file prepared by the author.