

Cosmological constraints from the angular power spectrum and bispectrum of luminous red galaxies and CMB lensing

Francesco Verdiani ^{a,b,c,d}, Léa Harscouet ^e, Matteo Zennaro ^e, David Alonso ^e
and Boryana Hadzhiyska ^{f,g}

^aScuola Internazionale Superiore di Studi Avanzati (SISSA),
Via Bonomea 265, 34136 Trieste, Italy

^bINFN Sezione di Trieste,
Via Valerio 2, 34127 Trieste, Italy

^cIstituto Nazionale di Astrofisica, Osservatorio Astronomico di Trieste,
Via Tiepolo 11, 34143 Trieste, Italy

^dInstitute for Fundamental Physics of the Universe,
Via Beirut 2, 34151 Trieste, Italy

^eDepartment of Physics, University of Oxford,
Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, United Kingdom

^fInstitute of Astronomy,
Madingley Road, Cambridge CB3 0HA, United Kingdom

^gKavli Institute for Cosmology Cambridge,
Madingley Road, Cambridge CB3 0HA, United Kingdom

E-mail: fverdian@sissa.it, lea.harscouet@physics.ox.ac.uk,
matteo.zennaro@physics.ox.ac.uk, david.alonso@physics.ox.ac.uk,
boryanah@ast.cam.ac.uk

ABSTRACT: We study the projected clustering of photometric luminous red galaxies from the DESI Legacy Survey, combining their angular power spectrum, bispectrum, and cross-correlation with maps of the CMB lensing convergence from the *Planck* satellite. We employ a perturbative bias expansion in Eulerian space to describe the clustering of galaxies, modelling the power spectrum and bispectrum at one-loop and tree level, respectively. This allows us to use the bispectrum to self-consistently calibrate the perturbative bias parameters. We validate this model against an N -body simulation, and show that it can be used up to scales of at least $k_{\max}^P \simeq 0.2 h \text{ Mpc}^{-1}$ and $k_{\max}^B \simeq 0.08 h \text{ Mpc}^{-1}$, saturating the information recovered from the data. We obtain constraints on the amplitude of matter fluctuations $\sigma_8 = 0.761 \pm 0.020$ and the non-relativistic matter fraction $\Omega_m = 0.307 \pm 0.015$, as well as the combination $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3} = 0.769 \pm 0.020$. Including the galaxy bispectrum leads to a 10–20%

improvement on the cosmological constraints, which are also in good agreement with previous analyses of the same data, and in mild tension with *Planck* at the $\sim 2.5\sigma$ level. Finally, using the bispectrum allows for a substantially more precise measurement of the bias parameters of this sample, which are in reasonable agreement with existing coevolution relations.

KEYWORDS: cosmological parameters from LSS, galaxy clustering, gravitational lensing, power spectrum

ARXIV EPRINT: [2510.17796](https://arxiv.org/abs/2510.17796)

Contents

1	Introduction	1
2	Data	3
2.1	Public galaxy and CMB lensing data	3
2.2	Power spectrum and bispectrum estimation	3
3	Methodology	5
3.1	Simulations	5
3.2	Theory model	6
3.3	Accelerating theory predictions	10
3.4	Likelihood	11
4	Results	12
4.1	Validation on simulations	13
4.2	Cosmological constraints from 2- and 3-point correlators	17
4.3	Galaxy bias relations from projected clustering	23
5	Conclusions	25
A	Additional tests	27
A.1	Treatment of redshift-space distortions	27
A.2	Simpler bias parametrisations	27
A.3	Emulators accuracy	29

1 Introduction

The projected clustering of galaxies is one of the oldest probes of the large-scale structure (LSS) in existence. Early measurements of the angular correlation function of galaxies and clusters of galaxies from optical and radio surveys were vital in establishing the concept of galaxy bias, providing evidence for the large-scale homogeneity of the Universe, and confirming the basic predictions of gravitationally-led structure formation [1–5].

The amount of cosmological information that can be extracted from the angular clustering of galaxies in photometric surveys is relatively limited by comparison with their spectroscopic counterparts, where information from the full three-dimensional distribution of galaxies can be exploited [6–9]. Nevertheless, photometric surveys have a number of advantages, most importantly the large number density of sources that can be photometrically identified. This allows us to study their clustering over smaller scales before reaching the shot noise-dominated regime, to increase our sensitivity to important effects, such as lensing magnification, and to obtain high-precision measurements of the cross-correlation between galaxies and other projected tracers of the LSS [10–15]. In particular, the combination of these cross-correlations with the auto-correlation of galaxies at different redshifts enables a *tomographic* reconstruction of the key physical quantities that these probes are sensitive

to [16–20]. Within this general framework, *lensing tomography*, the combination of galaxy clustering and their cross-correlation with measurements of the gravitational lensing effect, has now become an established and high-precision cosmological probe. Lensing tomography analyses have allowed us to reconstruct the universe’s growth history at high precision, reaching redshifts $z \sim 3$ [21–23], as well as measuring the cosmological parameters governing the distribution evolution of cosmic structures at late times, such as the matter fraction Ω_m , or the expansion rate H_0 .

Lensing tomography, as well as other data combinations employing the angular clustering of galaxies, are set to become more important with the advent of Stage-IV [24] imaging surveys, such as Euclid [25], the Rubin Observatory [26, 27], and the future Roman Space telescope [28], as well as surveys at other wavelengths. It is therefore important to ensure that as much information as possible is recovered from these observations, including any potentially new effect that could become observable given the increased sensitivity that these experiments will achieve. In particular, the study of higher-order correlators of the projected galaxy distribution, such as the bispectrum, has lagged behind the use of similar statistics in other probes, such as spectroscopic clustering [29–32], the Cosmic Microwave Background (CMB) [33, 34], and cosmic shear [35, 36]. In part this is due to the relative complexity inherent to bispectrum analysis. First, bispectrum estimators are typically significantly more computationally costly than power spectrum methods. Secondly, the resulting data vector is often much larger, involving a large number of triangle configurations. Thirdly, estimating the covariance matrix of the measurements using simulations can be prohibitively expensive. Finally, the regime over which different models for the galaxy bispectrum are sufficiently accurate has not been as thoroughly studied as the case of the power spectrum, and obtaining theoretical predictions is also often expensive. The first three problems in this list may be tackled through the use of novel bispectrum estimators [37–40], such as the “filter-square” approach first presented in [41]. Likewise, advances in machine learning-based emulators may be used to accelerate the theoretical calculations.

In this paper we carry out the first comprehensive analysis of the angular power spectrum and bispectrum of galaxies in combination with CMB lensing. Our work will build on the analysis of [42] (H25 hereafter), where the measurements of the galaxy power- and bispectrum used here, as well as their first rudimentary analysis, were presented. Crucially, we will use simulations to quantify the regime over which a perturbative treatment of the galaxy bias may be safely used in a joint analysis of 2- and 3-point correlators, before applying this framework to the real data. We will show that including the bispectrum allows us to self-consistently measure the free parameters of the bias expansion at high precision, thereby improving the final cosmological constraints, while demonstrating the internal consistency of the model.

This paper is structured as follows. Section 2 presents the datasets used, as well as the methods used to generate the power spectrum and bispectrum measurements used in our analysis. The model used to describe these measurements, the simulations used to validate this model, and the method used to obtain constraints on the model parameters are described in section 3. The results of our analysis, including both the validation on simulations and the constraints found on real data, are presented in section 4. We summarise our main results and conclude in section 5.

2 Data

2.1 Public galaxy and CMB lensing data

We use the publicly available *Planck* PR4 CMB lensing map, specifically the Generalized Minimum Variance (GMV) map, constructed from a joint Wiener filtering of the temperature and polarisation maps. The angular multipoles of the map are first rotated to Equatorial coordinates before being low-pass filtered with a top-hat filter that nulls all harmonic coefficients with $\ell > 3N_{\text{side}}$ and being transformed into a real-space map. Here $N_{\text{side}} = 2048$ is the HEALPix pixelisation resolution parameter used, corresponding to a pixel size of $\delta\theta \simeq 1.7'$. Similarly, the publicly available sky mask associated with this map is rotated to Equatorial coordinates in pixel space, and then apodised using a ‘‘C1’’ apodisation kernel with angular scale 0.2° , as implemented in NaMaster.¹ We then make the resulting mask binary, setting all values above 0.5 to 1, and all other values to zero. The impact of contamination from thermal Sunyaev-Zel’dovich and Cosmic Infrared Background foregrounds on CMB lensing cross-correlations with the *Planck* data were found to be largely negligible in [13, 43].

We also use galaxy overdensity maps constructed from the photometric luminous red galaxy (LRG) sample from the 9th data release (DR9) of the Legacy Survey [44]. Specifically, we use the galaxy overdensity maps and associated masks of the ‘‘main’’ sample made publicly available by [45], and used in the analysis of [13] (S25 hereafter), which reports the angular number densities along other properties of the sample in table 1. We do not perform any further post-processing steps on these maps, which are provided natively at our chosen resolution $N_{\text{side}} = 2048$. Maps are provided for four different redshift bins, spanning the redshift range $0.3 \lesssim z \lesssim 1.3$. The footprint of the LRGs is near-identical in all redshift bins, and its overlap with the CMB lensing footprint covers $\sim 42\%$ of the sky, which corresponds to about 17340 deg^2 . The redshift distributions of each bin are also publicly available, and shown for completeness in the left panel of figure 1. We use these redshifts to build our theoretical predictions for the angular power spectra and bispectra. The final ingredient needed for this analysis is an estimate of the slope in the cumulative number counts of sources as a function of magnitude s_μ , which we use to quantify the impact of magnification bias. We use the values of s_μ listed in table 1 of S25. For further details regarding the sample selection and characterisation, the derivation of contaminant weights, redshift distributions, and number count slopes, we refer readers to [45] and S25.

2.2 Power spectrum and bispectrum estimation

We estimate the galaxy-galaxy angular power spectrum C_ℓ^{gg} in each redshift bin, as well as its cross-correlation with CMB lensing convergence, $C_\ell^{g\kappa}$, using the pseudo- C_ℓ algorithm [PCL, 46] as implemented in NaMaster [47]. All power spectra were estimated using bandpowers of width $\Delta\ell = 9$, in the range $2 \leq \ell \leq 6140$. The impact of residual mode coupling due to the presence of sky masks is forward-modelled in the theoretical predictions using the bandpower window functions predicted by NaMaster. As per the procedure laid out in appendix C of S25, an additional correction must be performed on the galaxy-convergence cross-spectrum in order to account for different masking procedures and inhomogeneous

¹<https://github.com/LSSTDESC/NaMaster>

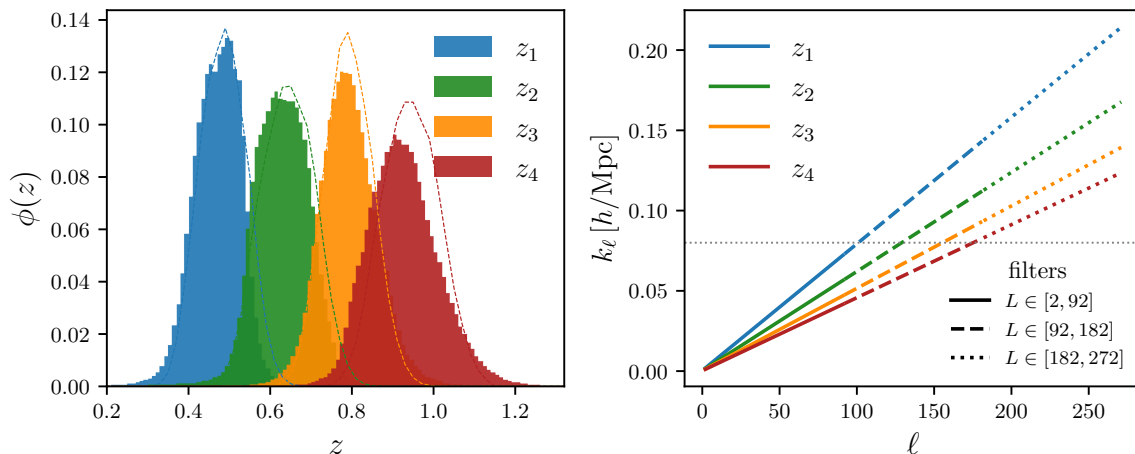


Figure 1. *Redshift distribution and scale cuts.* *Left:* Redshift distribution for each tomographic bin used here (filled histograms). The distribution of the simulated ABACUSSUMMIT sample is shown as dashed lines. *Right:* Multipole-wavenumber relation, in the Limber approximation, for the filtered bispectrum data. The gray horizontal line marks the fiducial scale cut used here for the F²B measurements.

normalization introduced in the making of the CMB lensing convergence map κ . These effects are corrected using a simulation-based transfer function, computed over 480 simulated CMB lensing reconstructions [12]. Once computed, we divide the $C_\ell^{g\kappa}$ measurements by the transfer function (whose inverse is given by eq. (C.1) in S25).

We also estimate the angular galaxy bispectrum in each redshift bin, using the filtered-squared bispectrum (F²B) estimator presented in [41]. In this framework, the bispectrum of a given field δ_g is calculated as the cross-correlation between δ_g and the square of a filtered version of itself $(\delta_{g,L})^2 \equiv (W_L \otimes \delta_g)^2$, where the filter W_L suppresses all modes outside a range of multipoles around the angular scale L . The F²B is thus defined as

$$\Phi_{LL\ell}^{ggg} \equiv \langle (\delta_{g,L})_{\ell m}^2 \delta_{g,\ell m}^* \rangle, \quad (2.1)$$

and is related to the reduced bispectrum $b_{\ell_1\ell_2\ell_3}^{ggg}$ via

$$\Phi_{LL\ell}^{ggg} = \sum_{\ell_1\ell_2\ell_3} K_{LL\ell}^{\ell_1\ell_2\ell_3} b_{\ell_1\ell_2\ell_3}^{ggg}, \quad (2.2)$$

where the F²B kernel $K_{LL\ell}^{\ell_1\ell_2\ell_3}$ is

$$K_{LL\ell_b}^{\ell_1\ell_2\ell_3} \equiv h_{\ell_1\ell_2\ell_3}^2 \frac{W_{\ell_1}^L W_{\ell_2}^L W_{\ell_3}^{\ell_b}}{N_{\ell_b}}, \quad (2.3)$$

where we have explicitly noted that we use bandpowers, labelled by ℓ_b , with N_{ℓ_b} the total number of modes in the bandpower ($N_{\ell_b} \equiv \sum_{\ell \in \ell_b} (2\ell + 1) W_\ell^{\ell_b}$). Finally, the prefactor $h_{\ell_1\ell_2\ell_3}^2$ is

$$h_{\ell_1\ell_2\ell_3}^2 \equiv \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2. \quad (2.4)$$

Note that the cross-correlation between $(\delta_{g,L})^2$ and δ_g is estimated using the pseudo- C_ℓ method to account for mask-related mode coupling. As shown in [41], the residual mode

coupling due to carrying out the filtering operation on an incomplete sky is negligibly small. Thus defined, the F²B therefore recovers the angular bispectrum for triangular configurations that are close to isosceles, although, as noted in [41], this is nevertheless able to capture a large fraction of all available triangles. The estimator can also be easily generalised to capture arbitrary triangle configurations, as well as correlations between different fields [42].

The filter W_L defined above can take a variety of forms; in this study, we define three equal-width top-hat filters which span the range $2 \leq \ell \leq 272$, each $\Delta L = 90$ wide. For each filter W_L , there is a corresponding F²B measurement over the range $2 \leq \ell \leq 6140$, binned in 682 bandpowers of width $\Delta \ell = 9$. The data vector therefore contains the gg and $g\kappa$ power spectra, as well as a few F²B measurements — the number of F²B included in the data vector will depend on the scale cuts we implement in the analysis. Indeed, a F²B whose filter selects scales beyond the regime where our theoretical predictions are reliable should be discarded; the right panel of figure 1 shows the redshift-dependent relation between Fourier modes (where we define theoretical scale cuts) and harmonic multipoles (where the filters W_L are defined) which we use to inform our filter selection in each redshift bin. For instance, for a scale cut at $k_{\max} \sim 0.08 h/\text{Mpc}$ (shown as the horizontal dotted line in figure 1), it is safe to include the first F²B measurement for the lowest redshift bin, and the first two measurements for the highest redshift bin.

The covariance of the data vector is estimated using the procedure first described in [41] — for a detailed account of the various terms entering the estimate of the covariance for cross-correlations, please refer to appendix C of H25. It is a fully analytical approximation of the covariance, easily estimated from the data itself in a model-independent way. Reinterpreting the bispectrum as a power spectrum of fields δ_g and $\delta_{g,L}^2$ allows us to reuse already-existing covariance estimation methods for power spectra [see e.g. 48]; in doing so (which means, in practice, simply implementing the Gaussian power spectrum covariance from NaMaster), we recover all Gaussian and non-Gaussian diagonal contributions of the F²B and F²B-PCL covariance. Additional non-negligible off-diagonal contributions, called N_{222} and N_{32} for the F²B auto-covariance and F²B-PCL cross-covariance respectively, can be derived analytically, and are rescaled by the inverse of the effective sky fraction (to account for mode loss) before being added to the diagonal “Gaussian” covariance. In this analysis, we only include the N_{222} term and neglect the N_{32} term. This is because the latter, which is considerably smaller in amplitude than N_{222} in the $[\Phi_{LL\ell}^{ggg}, C_\ell^{gg}]$ covariance block, and simply negligible for the $[\Phi_{LL\ell}^{ggg}, C_\ell^{g\kappa}]$ block (see appendix C of [42]), takes longer to compute as it contains a generalized F²B, $\Phi_{L\ell\ell}$, in which one of the legs entering the squared field is filtered on a range of scales corresponding to the bandpowers. As shown in [41], adding N_{32} to the N_{222} correction only marginally affects the F²B residual distribution.

3 Methodology

3.1 Simulations

In order to validate the theoretical model described below, and quantify the range of scales over which we may be able to use it safely, we make use of simulation-based mock observations mimicking the data we use here.

ABACUSSUMMIT is a suite of cosmological N -body simulations designed to meet and exceed the Cosmological Simulation Requirements of the DESI survey [49]. The simulations were run with ABACUS [50, 51], a high-accuracy, high-performance cosmological N -body code optimized for GPU architectures and large-volume runs, on the Summit supercomputer at the Oak Ridge Leadership Computing Facility. In this work, we use the halo light cone catalogs from the `huge` resolution box, which spans $7.5 h^{-1}$ Gpc on a side with 8640^3 particles, each with mass $5.7 \cdot 10^{10} h^{-1} M_{\odot}$ (corresponding to a particle mass 27 times larger than the `base` resolution). These catalogs include the past light cone geometry needed for mock survey generation, and allow us to cover the full DESI footprint with sufficient volume [52]. The simulations adopt the fiducial *Planck* 2018 cosmology ($\Omega_b h^2 = 0.02237$, $\Omega_c h^2 = 0.12$, $h = 0.6736$, $10^9 A_s = 2.0830$, $n_s = 0.9649$).

To mimic the DESI Extended photometric LRG sample from Data Release 10 (DR10) [45, 53], which is based on imaging from DECaLS on Blanco, MzLS on Mayall, and BASS on Bok, we generate galaxy mocks from the ABACUSSUMMIT halo light cones using the `AbacusHOD` package. This method populates halos with galaxies according to a halo occupation distribution (HOD), for which we adopt the best-fit parameters for the DESI 1% LRGs at $z = 0.5$ and $z = 0.8$ from [54], linearly interpolated between these redshifts to provide continuous evolution across $0.3 < z < 1.1$. The resulting catalog is then randomly downsampled to match the observed number densities in the four redshift bins of our target dataset. The galaxies are binned into tomographic bins matching roughly the redshift range and distribution of the target sample. For this, each galaxy is assigned a Gaussian photometric redshift error with standard deviation $\sigma_z = 0.024(1+z)$, and then binned according to the resulting photometric redshift using the same bin edges that define our target LRG sample. The magnitude of the Gaussian photo- z error used was selected to approximately reproduce the width and tails of the redshift distributions in the real data. The resulting redshift distributions are shown in the left panel of figure 1.

We construct full-sky CMB lensing convergence maps using the density shells of the ABACUSSUMMIT `huge` light cone. The matter distribution is projected into ~ 1000 shells of thickness $\sim 10 h^{-1}$ Mpc between $z = 0.1$ and $z = 2.3$, and mapped onto HEALPix pixels at $N_{\text{side}} = 16384$ (corresponding to $0.23'$ resolution). We compute the convergence field under the Born approximation, with sources placed at $z = 1089.3$ corresponding to the period of recombination.

3.2 Theory model

3.2.1 Angular power spectra and bispectra

In this work we make use of two projected tracers of the large-scale structure: The angular overdensity of LRGs δ_g in different redshift bins, and maps of the CMB lensing convergence κ .

The angular galaxy overdensity, accounting for the impact of lensing magnification, can be related to the three-dimensional matter and galaxy overdensities (Δ_g and Δ_m , respectively) via

$$\delta_g(\hat{\mathbf{n}}) = \int d\chi [q_g(\chi)\Delta_g(\chi\hat{\mathbf{n}}, z(\chi)) + q_\mu(\chi)\Delta_m(\chi\hat{\mathbf{n}}, z(\chi))], \quad (3.1)$$

where χ is the radial comoving distance, $\hat{\mathbf{n}}$ is a line-of-sight unit vector, and the radial kernels associated with intrinsic clustering (q_g) and magnification (q_μ) are, assuming a spatially

flat universe,

$$q_g(\chi) = H(z(\chi)) p(z(\chi)), \quad (3.2)$$

$$q_\mu(\chi) = \frac{3}{2} H_0^2 \Omega_m (1+z) \chi \int_{z(\chi)}^\infty dz' p(z') (5s(z') - 2) \left(1 - \frac{\chi}{\chi(z')}\right). \quad (3.3)$$

Here $p(z)$ is the redshift distribution of the galaxy sample, Ω_m and H_0 are the matter fraction and the current expansion rate, respectively, and $s(z)$ is the slope of the cumulative number count of sources as a function of magnitude at the magnitude limit [45]. Following S25, we will assume that s does not vary significantly within the redshift bin, and adopt the values presented in [43]. It is worth noting that, since for this sample $s > 0.4$, magnification leads to a *positive* contribution to the observed galaxy overdensity (i.e. regions of positive matter overdensity will contribute towards a positive δ_g).

The CMB lensing convergence traces the matter overdensities directly via, again valid only for a spatially flat universe,

$$\kappa(\hat{\mathbf{n}}) = \int d\chi q_\kappa(\chi) \Delta_m(\chi \hat{\mathbf{n}}, z(\chi)), \quad q_\kappa(\chi) \equiv \frac{3}{2} H_0^2 \Omega_m (1+z) \chi \left(1 - \frac{\chi}{\chi_{\text{LSS}}}\right), \quad (3.4)$$

where χ_{CMB} is the distance to the last-scattering surface at $z_{\text{CMB}} \simeq 1089$. Note that this expression, as well as the description of magnification used above, neglects a scale-dependent prefactor in harmonic space, $f_\ell \equiv \ell(\ell+1)/(\ell+1/2)^2$, which is negligibly different from 1 on the scales used here.

Within the Limber approximation adopted here, the angular power spectra of δ_g and κ can be related to the 3D power spectra of Δ_g and Δ_m via

$$C_\ell^{gg} = \int \frac{d\chi}{\chi^2} \left[q_g^2(\chi) P_{gg}(k_\ell, z(\chi)) + 2q_g(\chi) q_\mu(\chi) P_{gm}(k_\ell, z(\chi)) + q_\mu^2(\chi) P_{mm}(k_\ell, z(\chi)) \right], \quad (3.5)$$

$$C_\ell^{g\kappa} = \int \frac{d\chi}{\chi^2} q_\kappa(\chi) [q_g(\chi) P_{gm}(k_\ell, z(\chi)) + q_\mu(\chi) P_{mm}(k_\ell, z(\chi))], \quad (3.6)$$

where $P_{gm}(k, z)$ and $P_{gg}(k, z)$ are the 3D galaxy-matter and galaxy-galaxy power spectra, respectively, and $k_\ell \equiv (\ell+1/2)/\chi$. At this point it is worth noting that the impact of lensing magnification is negligibly small in the galaxy auto-correlation, but becomes relevant at the $\sim 10\%$ level in the CMB lensing cross-correlation, leading to parameter shifts of order $\sim 1\sigma$ if neglected.

For simplicity, the expressions above have assumed the so-called Limber approximation [55], which is a rather good approximation given the relatively wide redshift bins used here. Nevertheless, we will calculate all power spectrum templates described in section 3.3 without adopting this approximation on scales $\ell < 100$, where non-Limber effects are most relevant [13]. Another potentially important large-scale effect is redshift-space distortions (RSDs) [56]. In the context of projected clustering, RSDs cause a spatially-correlated migration of galaxies in and out of the selected redshift bin, leading to additional large-scale power [57]. We incorporated the impact of linear RSDs in the galaxy auto-correlation, where it is most relevant. As we show in appendix A.1, the impact of these large-scale effects is relatively small, and our results are consistent with those found after removing the scales that are most affected by them.

Finally, our analysis will incorporate measurements of the projected galaxy bispectrum $b_{\ell_1\ell_2\ell_3}^{ggg}$, related to the F²B estimator through eq. (2.2). The projected and three-dimensional bispectra are then related via the Limber integral

$$b_{\ell_1\ell_2\ell_3}^{ggg} = \int \frac{d\chi}{\chi^4} q_g^3(\chi) B_{ggg}(k_{\ell_1}, k_{\ell_2}, k_{\ell_3}, z(\chi)). \quad (3.7)$$

As mentioned above, the impact of magnification is negligible in the galaxy auto-spectrum and, for the same reason, we will ignore its contribution to the galaxy bispectrum. This is further justified by the significantly larger statistical uncertainties of the bispectrum measurements.

3.2.2 Shot noise

We add a shot noise contribution to the model galaxy auto-spectrum of the form

$$C_\ell^{gg,SN} = \frac{1 + \alpha_P}{\bar{n}}, \quad (3.8)$$

where \bar{n} is the angular number density of sources, and α_P is a free parameter characterising deviations from purely Poissonian shot noise. This parameter thus allows us to account for other stochastic contributions to the small-scale clustering of galaxies that are effectively uncorrelated (and hence leading to a white noise component) within the scales studied.

We will include a similar contribution to the projected bispectrum, of the form

$$b_{\ell_1\ell_2\ell_3}^{ggg,SN} = \left(\frac{1 + \alpha_P}{\bar{n}} \right)^2 + \frac{1 + \alpha_P}{\bar{n}} \left[C_{\ell_1}^{gg(1)} + C_{\ell_2}^{gg(1)} + C_{\ell_3}^{gg(1)} \right], \quad (3.9)$$

where $C_\ell^{gg(1)}$ is the tree-level galaxy power spectrum without shot noise. The two terms in eq. (3.9) stem from different contribution of the finite number density to a three-point correlator, and depend on different powers of \bar{n} . For the explicit computation that leads to eq. (3.9) for the projected bispectra, we refer the reader to appendix D of [41]. Note that our model assumes implicitly that the impact of non-Poissonian shot noise can be taken into account in both the 2- and 3-point correlators by simply scaling the effective number density of sources assuming Poisson statistics ($\bar{n}_{\text{eff}} = \bar{n}/(1 + \alpha_P)$). Although this is not the most general model describing stochastic bias components in galaxy clustering, we find this approximation to be sufficiently accurate in this analysis.

3.2.3 Three-dimensional clustering

The final ingredient needed in our model is a description of the three-dimensional clustering statistics of the galaxy and matter overdensities, namely $P_{gg}(k, z)$, $P_{gm}(k, z)$, and $B_{ggg}(k_1, k_2, k_3, z)$.

Here we will use a perturbative bias expansion in Eulerian space, within the so-called Effective Field Theory of LSS [58, 59]. Specifically, we describe the galaxy and matter overdensities up to third order as

$$\Delta_m = \delta^{(1)} + \delta^{(2)} + \delta^{(3)}, \quad \Delta_g = b_1 \Delta_m + \frac{b_2}{2} (\delta^{(1)})^2 + \frac{b_s}{2} s^2 + \frac{b_{3nl}}{2} \psi_{nl} + \frac{b_{k^2}}{2} \nabla^2 \delta^{(1)}, \quad (3.10)$$

where $\delta^{(1)}$ is the linear matter overdensity, and $\delta^{(n)}$ is the contribution to the matter overdensity at n -th order in standard perturbation theory. The squared traceless tidal field s^2 is defined as

$$s^2 \equiv s_{ij}s^{ij}, \quad s_{ij} = (\partial_i\partial_j\nabla^{-2} - \delta_{ij}^K\delta^{(1)}), \quad (3.11)$$

and the third-order operator ψ_{nl} is defined in e.g. [60].

We model the galaxy-galaxy and galaxy-matter power spectra at one-loop order in standard perturbation theory (SPT), including EFT counterterm corrections:

$$\begin{aligned} P_{gg}(k) = & b_1^2 P_{1\text{-loop}}(k) + b_1 b_2 P_{\delta^{(2)}(\delta^{(1)})^2}(k) + b_1 b_s P_{\delta^{(2)}s^2}(k) + b_1 b_{3nl} P_{\delta^{(1)}\psi_{nl}}(k) \\ & + \frac{b_2^2}{4} P_{(\delta^{(1)})^2(\delta^{(1)})^2}(k) + \frac{b_s^2}{4} P_{s^2s^2}(k) + \frac{b_2 b_s}{2} P_{s^2(\delta^{(1)})^2}(k) + b_1 b_{k^2} P_{\delta^{(1)}\nabla^2\delta^{(1)}}(k) \end{aligned} \quad (3.12a)$$

$$\begin{aligned} P_{gm}(k) = & b_1 P_{1\text{-loop}}(k) + \frac{b_2}{2} P_{\delta^{(2)}(\delta^{(1)})^2}(k) + \frac{b_s}{2} P_{\delta^{(2)}s^2}(k) + \frac{b_{3nl}}{2} P_{\delta^{(1)}\psi_{nl}}(k) \\ & + \left(\frac{b_{k^2}}{2} + b_1 c_m \right) P_{\delta^{(1)}\nabla^2\delta^{(1)}}(k) \end{aligned} \quad (3.12b)$$

$$P_{mm}(k) = P_{1\text{-loop}}(k) + 2c_m P_{\delta^{(1)}\nabla^2\delta^{(1)}}(k). \quad (3.12c)$$

Here $P_{1\text{-loop}}(k)$ is the 1-loop SPT matter power spectrum, and all other power spectra of the form $P_{\mathcal{O}_1\mathcal{O}_2}(k)$ are the power spectra between operators \mathcal{O}_1 and \mathcal{O}_2 at lowest non-zero order in perturbation theory. For combinations involving the higher derivatives term, we only retain the leading order contribution given by $\langle\delta^{(1)}\nabla^2\delta^{(1)}\rangle$, ignoring the contributions from $\langle(\delta^{(1)})^2\nabla^2\delta^{(1)}\rangle$, $\langle s^2\nabla^2\delta^{(1)}\rangle$ and $\langle\nabla^2\delta^{(1)}\nabla^2\delta^{(1)}\rangle$ [61]. Note that we have added a counterterm for the matter field proportional to k^2 , with a free parameter c_m . A similar counterterm would also be present for the galaxy overdensity, although its amplitude is fully degenerate with the non-local bias parameter b_{k^2} .

In turn, we model the galaxy bispectrum at tree level in SPT, in which case:

$$B_{ggg}(k_1, k_2, k_3) = 2P(k_1)P(k_2) \left[b_1^3 F_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{b_1^2 b_2}{2} Q_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{b_1^2 b_s}{2} T_2(\mathbf{k}_1, \mathbf{k}_2) \right] + (\text{c.c.}), \quad (3.13)$$

where ‘‘c.c.’’ stands for the other two cyclic permutations of the three wavevectors, $P(k) \equiv P_{\delta^{(1)}\delta^{(1)}}(k)$ is the linear matter power spectrum, $F_2(\mathbf{k}_1, \mathbf{k}_2)$ is the second-order density SPT kernel, and Q_2 and T_2 are the corresponding kernels for the quadratic and tidal fields:

$$F_2(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{5}{7} + \frac{1}{2}\mu_{12} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7}\mu_{12}^2, \quad Q_2(\mathbf{k}_1, \mathbf{k}_2) \equiv 1, \quad T_2(\mathbf{k}_1, \mathbf{k}_2) \equiv \mu_{12}^2 - \frac{1}{3}, \quad (3.14)$$

with $\mu_{12} \equiv \mathbf{k}_1 \cdot \mathbf{k}_2 / k_1 k_2$. Note that we do not include non-local or higher-derivative contributions to the bispectrum, since these are mostly relevant in the presence of non-linear redshift-space distortions and at one-loop order [62].

As in S25, we neglect the redshift evolution of nuisance parameters in the Limber integrals, and evaluate them at the effective redshift defined by the galaxy clustering kernel

$$z_{\text{eff}} = \frac{\int (d\chi/\chi^2) q_g^2(\chi) z(\chi)}{\int (d\chi/\chi^2) q_g^2(\chi)} \quad (3.15)$$

In principle the other combinations of kernels, such as $q_g q_\kappa$ and q_g^3 , would entail a slightly different effective value. However, given the mild redshift evolution of the nuisance parameters observed in our results, we verified that this effect is negligible.

3.3 Accelerating theory predictions

Although the calculation of the theoretical predictions for the angular power spectra can be carried out with relative efficiency, particularly using fast schemes to calculate PT convolutions, such as FAST-PT [63, 64], the time to complete a fully converged Markov chain (see section 3.4) can still be significant (e.g. hours or days). This problem is exacerbated in the case of the angular bispectrum, given the large number of Limber integrals needed to cover all required triangle configurations, and the additional complication of convolving the result with the F²B kernel (eq. (2.3)). In order to increase the efficiency of our theory predictions as a function of cosmological parameters, we employ machine-learning based emulators. Concretely, we will emulate projected quantities, integrated over the radial kernels of the tracers under study. Thus, our emulator is strictly tied to the specific dataset considered in this work.

In both the 2-point and 3-point observables studied here, the bias coefficients and counterterms ($b_1, b_2, b_s, b_{3nl}, b_{k^2}, c_m$) are always multiplicative and can be brought outside the Limber integrals. For instance, consider the galaxy-convergence power spectrum in eq. (3.6). Within the perturbative bias expansion used here, it can be written as

$$C_\ell^{g\kappa} = b_1 C_{1\text{-loop},\ell}^{g\kappa} + \frac{b_2}{2} C_{\delta^{(2)}(\delta^{(1)})^2,\ell}^{g\kappa} + \frac{b_s}{2} C_{\delta^{(2)}s^2,\ell}^{g\kappa} + \frac{b_{3nl}}{2} C_{\delta^{(1)}\psi_{nl},\ell}^{g\kappa} + \left(\frac{b_{k^2}}{2} + b_1 c_m\right) C_{\delta^{(1)}\nabla^2\delta^{(1)},\ell}^{g\kappa} + C_{1\text{-loop},\ell}^{\mu\kappa} + c_m C_{\delta^{(1)}\nabla^2\delta^{(1)},\ell}^{\mu\kappa}, \quad (3.16)$$

where we have defined the angular power spectrum templates corresponding to the different PT power spectra in eq. (3.12):

$$C_{\mathcal{O}_1\mathcal{O}_2,\ell}^{ab} \equiv \int \frac{d\chi}{\chi^2} q_a(\chi) q_b(\chi) P_{\mathcal{O}_1\mathcal{O}_2}(k_\ell, z(\chi)). \quad (3.17)$$

A similar expression can be derived for C_ℓ^{gg} , with the addition of the shot noise component. Crucially, these templates depend only on a small number of cosmological parameters (two in our case, see below), and can therefore be emulated a priori.

In the same way for the galaxy bispectrum, where the acceleration enabled by emulators is even more essential, we may write

$$\Phi_{LLL}^{ggg} = b_1^3 F_{LLL}^{ggg} + \frac{b_1^2 b_2}{2} Q_{LLL}^{ggg} + \frac{b_1^2 b_s}{2} T_{LLL}^{ggg} + b_1^2 (1 + \alpha_P) \text{SN}_{2,LLL}^{ggg} + (1 + \alpha_P)^2 \text{SN}_{1,LLL}^{ggg}, \quad (3.18)$$

where we have defined the F²B templates:

$$X_{LLL}^{ggg} \equiv \sum_{\ell_1 \ell_2 \ell_3} K_{LLL}^{\ell_1 \ell_2 \ell_3} \mathbf{b}_{\ell_1 \ell_2 \ell_3}^{ggg,X}, \quad (3.19)$$

with the F²B kernel defined in eq. (2.3), and

$$\mathbf{b}_{\ell_1 \ell_2 \ell_3}^{ggg,X} \equiv \int \frac{d\chi}{\chi^4} q_g^3(\chi) [2P(k_{\ell_1})P(k_{\ell_2})X(\mathbf{k}_{\ell_1}, \mathbf{k}_{\ell_2}) + (\text{c.c.})] \quad \text{for } X \in \{F, Q, T\}, \quad (3.20)$$

$$\mathbf{b}_{\ell_1 \ell_2 \ell_3}^{ggg,\text{SN}_2} \equiv \frac{1}{\bar{n}} \left[C_{\delta^{(1)}\delta^{(1)},\ell_1}^{gg} + C_{\delta^{(1)}\delta^{(1)},\ell_2}^{gg} + C_{\delta^{(1)}\delta^{(1)},\ell_3}^{gg} \right], \quad \mathbf{b}_{\ell_1 \ell_2 \ell_3}^{ggg,\text{SN}_1} \equiv \frac{1}{\bar{n}^2} \quad (3.21)$$

As before, the X templates depend only on cosmological parameters.

We build one emulator per template and per redshift bin. The architecture of each of our emulators is a Feed-Forward Neural Network, with 2 hidden layers of 128 nodes each. The input layer of the network has size 2, corresponding to $\{\log(10^{10} A_s), \Omega_c h^2\}$. The output layer has the same size as our ℓ vector. For each emulator, we create a sample of 10,000 input-output pairs, which we split into a training set (holding 90% of the sample) and a test set (with the remaining 10%). We build and train our models with `Keras3` [65] and `TensorFlow` [66], using a Rectified Linear Unit (ReLU) activation function, and optimizing our weights and biases with the Adam gradient descent algorithm [67], with exponentially decreasing learning rate. We use mean squared errors between the training and test sets as a loss function and, during training, also compute an accuracy metric that accounts for the percentage of output nodes differing less than 0.5% from the ground truth. We use this metric to establish an early stopping criterion. The emulators we obtain with this procedure are able to reproduce the corresponding test sets at better than 1% accuracy on all of the considered scales, see appendix A.3.

The numerical computation of the templates appearing in eq. (3.20), with the method developed in H25, takes ≈ 5 s when already emulating the linear power spectrum [68], rather than calling the Einstein-Boltzmann solver, and interpolating the evaluation of eq. (3.19). Still, this makes posterior sampling, which we find requires evaluating the likelihood $\mathcal{O}(10,000\text{--}100,000)$ times, practically unfeasible, especially when combining the redshift bins. Our approach of emulating the templates reduces the single evaluation time of eq. (3.20) by three orders of magnitude, down to ≈ 5 ms, enabling a feasible, and extremely fast, exploration of the posterior with MCMC methods.

3.4 Likelihood

We derive parameter constraints from the measurements described in section 2 assuming a Gaussian likelihood of the form

$$-2 \log p(\mathbf{d}|\vec{\theta}) = (\mathbf{d} - \mathbf{t}(\vec{\theta}))^T \mathbf{C}^{-1} (\mathbf{d} - \mathbf{t}(\vec{\theta})) + K, \quad (3.22)$$

where \mathbf{d} is the data vector, $\mathbf{t}(\vec{\theta})$ is the theory prediction, described in the preceding sections and dependent upon a set of free parameters $\vec{\theta}$, and K is an arbitrary normalisation constant.

In our analysis, \mathbf{d} will comprise a set of power spectrum (C_ℓ^{gg} and/or $C_\ell^{g\kappa}$) and F²B measurements ($\Phi_{LL\ell}^{ggg}$), with the covariance matrix estimated as described in section 2.2. The range of scales over which different power spectra and bispectra are used was determined following the validation against simulations presented in section 4.1. In our fiducial analysis we will use power spectrum data on angular scales $\ell < \ell_{\max}^P \equiv k_{\max}^P \chi(\bar{z})$, where $\chi(\bar{z})$ is the comoving distance to the mean redshift of the sample, and $k_{\max}^P = 0.2 h \text{ Mpc}^{-1}$. Likewise, we will include F²B measurements on scales $(L_{\text{mean}}, \ell) < \ell_{\max}^B \equiv k_{\max}^B \chi(\bar{z})$, with $k_{\max}^B = 0.08 h \text{ Mpc}^{-1}$. Here L_{mean} is the mean multipole of the top-hat filter defined in section 2.2; our fiducial criterion thus corresponds to including the first filter for all bins, as well as the second filter for z_3 and z_4 . Note that we find that the constraints on Ω_m and σ_8 do not improve significantly by including smaller scales. In addition to this, we remove the largest angular scales in the power spectrum measurements, with $\ell < 29$ (i.e. we discard the first three bandpowers).

Parameter	Prior	Parameter	Prior
$\Omega_c h^2$	$\mathcal{U}(0.1, 0.14)$	b_{3nl}	$\mathcal{U}(-10, 10)$
$\log(10^{10} A_s)$	$\mathcal{U}(2.3, 3.5)$	b_{k^2}	$\mathcal{N}(0, 15 \text{ Mpc}^2)$
b_1	$\mathcal{U}(1.2, 3.2)$	c_m	$\mathcal{N}(0, 10 \text{ Mpc}^2)$
b_2	$\mathcal{U}(-5, 5)$	s_μ	$\mathcal{N}(s_\mu^{\text{fid}}, 0.1)$
b_s	$\mathcal{U}(-10, 10)$	α_P	$\mathcal{N}(0, 0.3)$

Table 1. Model parameters and priors used.

The free parameters of our model will be the full set of bias expansion parameters, including the counter-term and shot-noise amplitude, $\{b_1, b_2, b_s, b_{k^2}, c_m, \alpha_P\}$, as well as two cosmological parameters, the physical abundance of cold dark matter $\Omega_c h^2$, and the amplitude of primordial scalar fluctuations $\log(10^{10} A_s)$. We fix all other Λ CDM cosmological parameters as follows: Following [13], we use $\Omega_m h^3 = 0.09633$ [69] as a proxy for the angular acoustic scale, measured by *Planck* at very high precision, as well as $\Omega_b h^2 = 0.02236$, $n_s = 0.9649$ (also measured by *Planck*). We also fix the sum of neutrino masses to $M_\nu = 0.06$ eV. Although the parameters we will sample are $\Omega_c h^2$ and $\log(10^{10} A_s)$, we will most commonly present our constraints in terms of the matter fraction Ω_m and the linear matter power spectrum amplitude σ_8 , as derived parameters. When considering multiple redshift bins in the same dataset, we will assume that the clustering of galaxies and matter in each bin is described by a different set of bias parameters (including α_P and c_m).

We sample the posterior distribution using *nautilus*, a nested sampling algorithm that optimises the way its live points explore the parameter space thanks to the use of Neural Networks [70]. We use 1000 live points, with 4 Neural Networks, and with convergence criterion $f_{\text{live}} = 0.01$ and $n_{\text{eff}} = 5000$. To construct this posterior we impose the priors on our model parameters listed in table 1. The flat priors on $\Omega_c h^2$ and $\log(10^{10} A_s)$ coincide with the parameter ranges over which the emulator described in section 3.3 is trained, and were chosen to ensure that they enclose the region of parameter space that can be constrained by the data in each redshift bin. The priors for b_2 , b_s , and b_{3nl} were chosen to be sufficiently broad to encompass the ranges over which they may be constrained by two-point function data alone. The prior on b_{k^2} mimics the prior used in [13] for the counterterm amplitude and the prior on c_m is based on the arguments described in [71], whereby

$$c_m \simeq \mathcal{O}(1) \times \frac{1}{k_{\text{NL}}^2} \approx \mathcal{O}(1) \times 5 h^{-2} \text{ Mpc}^2 \quad (3.23)$$

for $k_{\text{NL}}(z = 0.5) \simeq 0.45 h \text{ Mpc}^{-1}$.

4 Results

We now present the results of our analysis, beginning with a validation of the model using simulations before applying it to the real LRG data.

4.1 Validation on simulations

Simulation data provides a controlled setting, where the true underlying cosmology is known, to assess the performance of the theoretical model employed.

A key aspect of full-shape fits of galaxy clustering data is understanding the regime of validity of the model, i.e. defining respectively the largest and smallest scales that can be employed for the purpose of parameter inference. While the limiting factor on the largest scales is usually the presence of systematics, or the potential mismodelling of large-angle effects (e.g. the Limber approximation, redshift-space distortions, etc.), the modelling on smaller scales is limited by non-linearities in the clustering. The latter is particularly subtle to assess, and several approaches have been proposed in the literature, most of which rely on *a posteriori* validation by varying the small-scale cut k_{max} in the fit. In general, we expect the fully theoretical, first-principle, model based on 1-loop (tree level) EPT to be able to describe the mildly non-linear scales in the real-space halo and galaxy power spectrum up to a k_{max}^P of order $\approx 0.3 h \text{ Mpc}^{-1}$ ($\approx 0.1 h \text{ Mpc}^{-1}$), as extensively validated for 3D clustering on N-body simulations [72–74] for the redshifts of interest to this work.

With this aim, we thus start by analysing the mock LRG catalogue and CMB lensing map generated from the ABACUSSUMMIT simulation as described in section 3.1. We analyse the simulated data using the same pipeline as for the real data, measuring the galaxy auto-spectrum, the galaxy- κ cross-spectrum, and the ggg F²B. We will then extract parameter constraints from these measurements using the covariance matrix estimated for the real data, allowing us to quantify the amplitude of any potential parameter biases against their expected measurement uncertainties in the real data. Note, however, that the simulated data vector has intrinsically smaller statistical uncertainties than the data, since the catalogue and κ map cover the full celestial sphere, and the κ map does not include any noise contribution from lensing reconstruction. This means that we should expect the scatter in the best-fit parameters extracted from the simulation to be much smaller than the posterior uncertainties, so any significant deviation from the true cosmology of the simulation will be likely due to the limitations of the theory model used.

To quantify the adequacy of the theoretical model, we consider two key performance metrics commonly employed in the literature [75]. Specifically, we compute the figure of Bias (FoB), which quantifies the unbiasedness of the model in the recovery of the fiducial parameters, and the figure of Merit (FoM), which measures statistical power in constraining the parameters. These are computed for the two cosmological parameters $\vec{\theta} = \{\Omega_m, \sigma_8\}$ as a function of the small-scale cut. The FoB and FoM are defined as

$$\text{FoB} \equiv \sqrt{(\langle \vec{\theta} \rangle - \vec{\theta}_{\text{true}})^T \text{Cov}_{\vec{\theta}}^{-1} (\langle \vec{\theta} \rangle - \vec{\theta}_{\text{true}})}, \quad \text{FoM} \equiv \frac{1}{\sqrt{\det(\text{Cov}_{\vec{\theta}})}}, \quad (4.1)$$

where $\langle \vec{\theta} \rangle$ is the mean of the posterior distribution for the model parameters $\vec{\theta}$, $\vec{\theta}_{\text{true}}$ is the true value of the parameter θ , and $\text{Cov}_{\vec{\theta}}$ is the covariance of the posterior. The FoB thus quantifies the bias induced on the input parameters as a fraction of their uncertainties, while FoM quantifies the constraining power of the model.

We begin by quantifying the scales over which the one-loop bias model used here is able to describe the projected power spectra C_{ℓ}^{gg} and $C_{\ell}^{g\kappa}$. For this, we analyse the simulated

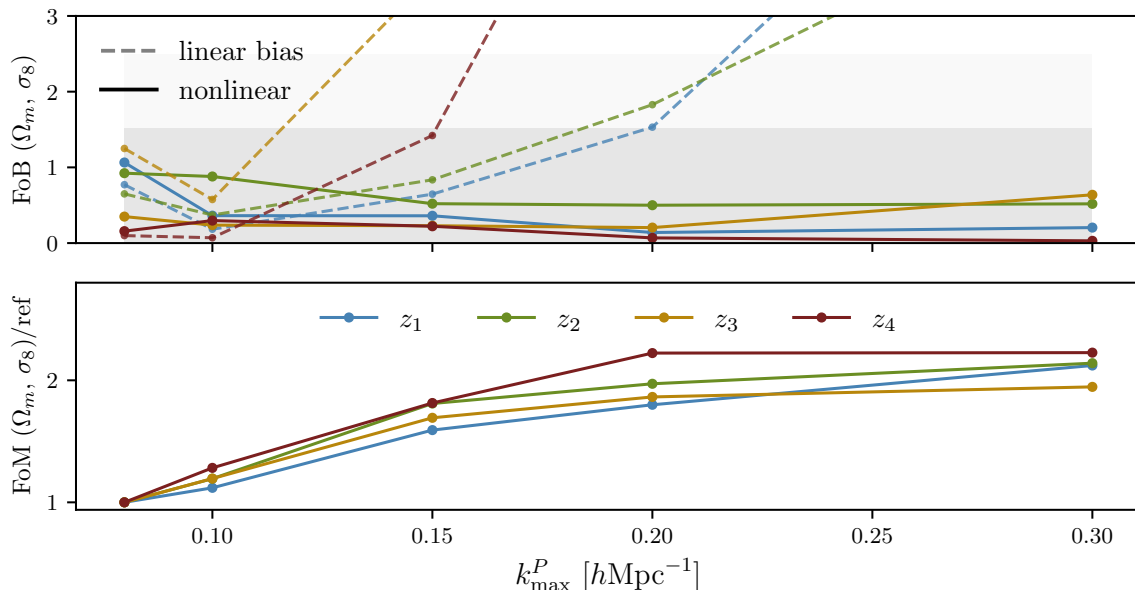


Figure 2. *Model performance in simulated data.* Figure of Bias (FoB, top panel), and figure of Merit (FoM, bottom panel) for the galaxy-galaxy and galaxy- κ angular power spectra in each redshift bin, for the two cosmological parameters $\{\Omega_m, \sigma_8\}$, as a function of the small-scale cut k_{\max}^P . Note that the FoM is reported relative to its value for $k_{\max}^P = 0.08 h \text{ Mpc}^{-1}$, in order to show results from all bins on the same scale. The solid lines refer to the nonlinear bias model assumed in our fiducial analysis, while dashed lines refer to a simplistic linear bias model. While linear bias can only give unbiased cosmological results up to $k_{\max}^P = 0.1 h \text{ Mpc}^{-1}$, the nonlinear model can extend to $k_{\max}^P = 0.3 h \text{ Mpc}^{-1}$ (but without significant gains in terms of extracted information for $k_{\max}^P > 0.2 h \text{ Mpc}^{-1}$).

measurements excluding the F²B for now. Figure 2 reports the result of the fits for each individual redshift bin. For reference, we also report the FoB for a “linear bias” model in which b_1 is the only nuisance parameter. This simpler model, while recovering unbiased results at very large scales ($k_{\max}^P \lesssim 0.1 h \text{ Mpc}^{-1}$), fails shortly afterwards. The apparently counter-intuitive fact that the metric, at a given scale cut, is worse for higher-redshift bins is an artifact of plotting as a function of comoving scales (as opposed to angular scales): A given k_{\max}^P corresponds to a higher multipole at higher redshifts, and the measurements thus achieve a stronger constraining power (i.e. smaller uncertainties). In contrast, the higher-order bias model used here (see section 3) consistently returns unbiased constraints on the two cosmological parameters for the various scale cuts and redshift bins that we consider. Furthermore, while there is a notable increase in constraining power when including mildly non-linear scales, this seemingly saturates already at $k_{\max}^P \sim 0.3 h \text{ Mpc}^{-1}$, or earlier ($k_{\max}^P \sim 0.2 h \text{ Mpc}^{-1}$) for the higher-redshift bins. This saturation can be related to the impact of shot noise and CMB lensing reconstruction noise on the small-scale statistical uncertainties, as well as the model flexibility allowed by the multiple bias parameters in the model, particularly the counter-term contributions proportional to $k^2 P(k)$. This is explored further in appendix A.2. Given these results, we make the conservative choice of setting the same small-scale cut of $k_{\max}^P = 0.2 h \text{ Mpc}^{-1}$ for all the redshift bins, which we apply to both C_ℓ^{gg} and $C_\ell^{g\kappa}$.

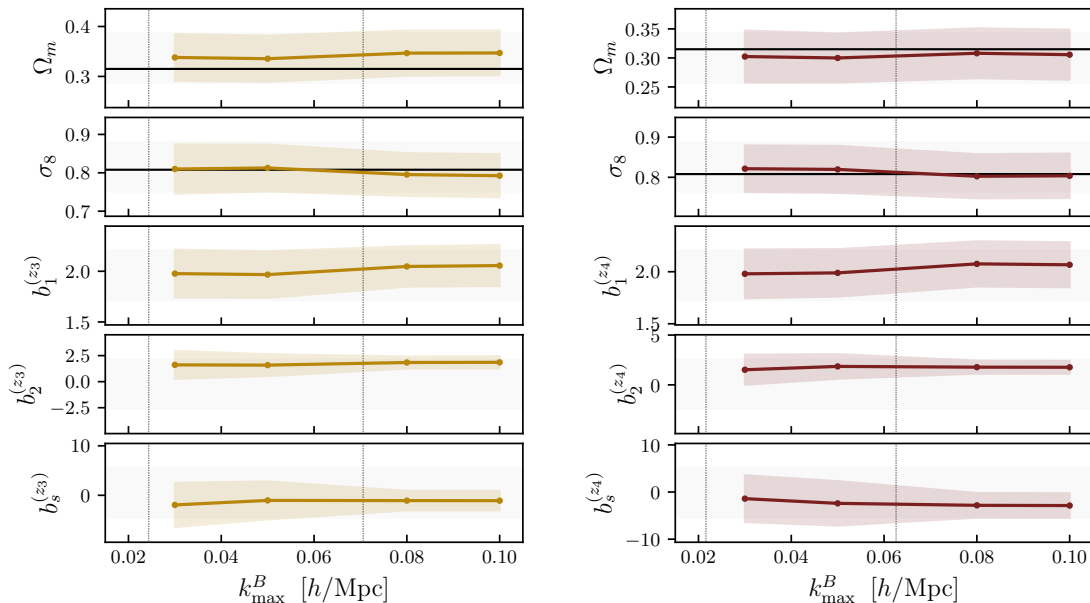


Figure 3. *Consistency with bispectrum scale cut on simulation data.* Posterior constraints on the model parameters as a function of the small-scale cut used for the galaxy bispectrum k_{\max}^B . Results are shown for two redshift bins, z_3 and z_4 (left and right panels, respectively). Similar results are found for the first two redshift bins. The shaded coloured bands show the 68% constraints on each parameter using the data covariance, with the gray band showing the constraints found when using only power spectrum data with $k_{\max}^P = 0.08 h \text{ Mpc}^{-1}$ (and $\ell_{\min} = 29$). Vertical lines represent, incrementally, the high- ℓ bound of the different filters used.

To estimate a reliable scale cut for the three-point correlation function, we carry out a separate, dedicated analysis. The main difficulty in doing so is the fact that, on its own, $\Phi_{LL\ell}^{ggg}$ is not able to place meaningful constraints on cosmological and bias parameters simultaneously. For the same reason, the same metrics used above (FoB and FoM), if applied to the combination of C_{ℓ}^{gg} , $C_{\ell}^{g\kappa}$ and $\Phi_{LL\ell}^{ggg}$, would be dominated by the two-point functions, making them almost insensitive to the bispectrum. Thus, we adopt a different strategy, combining $\Phi_{LL\ell}^{ggg}$ with C_{ℓ}^{gg} and $C_{\ell}^{g\kappa}$, but including the latter two only on the largest scales ($k_{\max}^P = 0.08 h \text{ Mpc}^{-1}$), thus minimising the information extracted by the two-point function. Furthermore, rather than simply examining the FoB and FoM directly, we will instead quantify the stability of the posterior constraints on all parameters as a function of k_{\max}^B . Figure 3 shows the constraints on all cosmological and bias parameters for the two highest redshift bins as a function of k_{\max}^B in this setup. Similar results were found for the other two redshift bins. We find that the preferred values of both Ω_m and σ_8 are largely stable with respect to the choice of k_{\max}^B , and their statistical uncertainties only vary slowly within the range of scales studied. In turn, the constraints on the bias parameters benefit strongly from the inclusion of the three-point function. Their preferred value is also stable with respect to k_{\max}^B . Given these results, we choose a scale cut of $k_{\max}^B = 0.08 h \text{ Mpc}^{-1}$, compatible with the values assumed in three-dimensional clustering studies [71, 76], as a compromise between the final size of the data vector and the desire to improve the constraints on the higher-order bias parameters.

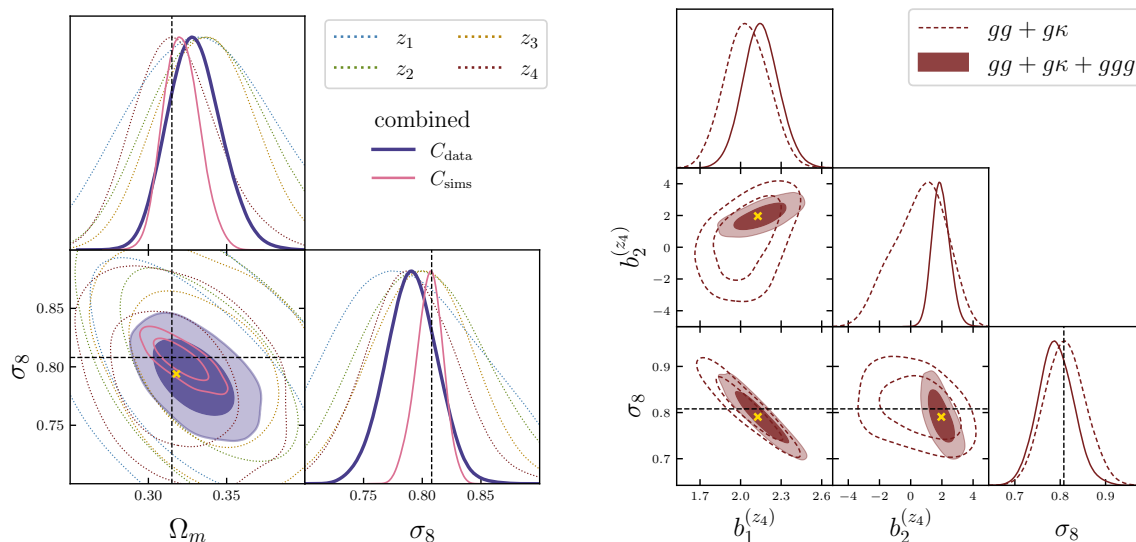


Figure 4. Constraints from the ABACUSSUMMIT simulated data. *Left:* Posterior, for single and combined redshift bins, of the cosmological parameters in the combined $gg + g\kappa + ggg$ analysis. The dashed lines show the true cosmological parameters of the ABACUSSUMMIT simulation, while the yellow cross marks the best-fit parameters found using the data covariance in the likelihood. The solid purple contours show the constraints obtained from the combination of all redshift bins, with the dotted contours showing the constraints found from individual redshift bins, all using the data covariance. The constraints found using the simulation covariance (significantly smaller due to the larger sky area and lack of lensing reconstruction noise) are shown in pink. The true cosmological parameters are recovered in all cases with no significant bias. *Right:* Posterior, for the highest redshift bin, showing the relevant degeneracies of σ_8 with the bias parameters with and without the 3-point information, using the data covariance. This shows how the 1D posterior for σ_8 is affected by the marginalization over the bias parameters, especially along the degeneracy with b_2 , when new information is added, leading also to a small shift in the central value.

Having chosen our fiducial scale cuts, $(k_{\text{max}}^P, k_{\text{max}}^B) = (0.2, 0.08) h \text{Mpc}^{-1}$, we now derive constraints from the full data vector, combining all two- and three-point correlations from the four redshift bins. The resulting constraints on Ω_m and σ_8 are shown in the left panel of figure 4. The constraints found using the data covariance matrix are compatible with the true cosmology of the ABACUSSUMMIT simulations within 1σ , and the best-fit parameter values (defined as the point in the MCMC chain with the lowest χ^2) are very close to the truth. Similar results are found for each redshift bin when analysed individually, and these are also shown in the figure for comparison. To determine whether the small shift observed between the posterior constraints and the true parameter values is a consequence of volume/projection effects or a genuine parameter bias, we repeat our analysis using the covariance matrix of the simulation (which, as discussed above, leads to significantly smaller data errors). The parameter constraints found in this case, shown in figure 4, are significantly tighter and perfectly compatible with the true parameter values. This suggests that the small parameter shift observed with the data covariance is indeed a result of projection effects. To confirm this further, the right panel of figure 4 shows the constraints on σ_8 and the b_1 and b_2 bias parameters of the fourth redshift bins. We can see that the multi-parameter

z_{bin}	Dataset	Ω_m	σ_8	fit χ^2 (PTE)
z_1	$gg + g\kappa$	0.296 ± 0.034	0.686 ± 0.058	1.07 (0.35)
	$gg + g\kappa + ggg$	0.294 ± 0.033	0.686 ± 0.055	0.98 (0.50)
z_2	$gg + g\kappa$	0.329 ± 0.035	0.772 ± 0.050	0.96 (0.55)
	$gg + g\kappa + ggg$	0.337 ± 0.036	0.761 ± 0.046	0.91 (0.69)
z_3	$gg + g\kappa$	0.297 ± 0.025	0.773 ± 0.042	1.02 (0.44)
	$gg + g\kappa + ggg$	0.302 ± 0.024	0.758 ± 0.036	0.99 (0.52)
z_4	$gg + g\kappa$	0.290 ± 0.023	0.785 ± 0.040	0.93 (0.65)
	$gg + g\kappa + ggg$	0.296 ± 0.022	0.765 ± 0.035	0.89 (0.78)
$z_2 + z_3 + z_4$	$gg + g\kappa$	0.301 ± 0.016	0.774 ± 0.023	0.96 (0.64)
	$gg + g\kappa + ggg$	0.307 ± 0.015	0.761 ± 0.020	0.96 (0.65)
All bins	$gg + g\kappa$	0.299 ± 0.014	0.759 ± 0.020	0.99 (0.54)
	$gg + g\kappa + ggg$	0.303 ± 0.014	0.753 ± 0.018	0.97 (0.63)
All bins $\{b_1, b_2, b_s\}$	$gg + g\kappa$	0.299 ± 0.011	0.778 ± 0.020	0.95 (0.72)
	$gg + g\kappa + ggg$	0.300 ± 0.010	0.764 ± 0.018	0.95 (0.75)
All bins $\{b_1, b_{k^2}, c_m\}$	$gg + g\kappa$	0.295 ± 0.011	0.772 ± 0.021	0.94 (0.74)
	$gg + g\kappa + ggg$	0.293 ± 0.009	0.785 ± 0.016	0.98 (0.61)

Table 2. *Cosmological parameter constraints* found from different data combinations. Results are shown for each redshift bin independently, as well as for the combination of the last three bins and all four bins. The last two sets of rows show the constraints found when analysing the data under the simplified bias models explored in appendix A.2. In each case, we show constraints from both two-point and two- and three-point correlators. The last column shows the reduced χ^2 and probability-to-exceed for the best-fit model.

posterior distribution is clearly non-Gaussian, and marginalisation over the bias parameters, particularly in the case of b_2 , can lead to a shift in the marginalised posterior distribution. We also verified that with the true simulation covariance this shift moves the posterior closer to the true cosmology, the $gg + g\kappa + ggg$ result being shown in pink in the left panel of figure 4.

In conclusion, we find that the bias model used here can describe our measurements with sufficient accuracy within our fiducial scale cuts. In appendix A.2 we explored the ability of simpler bias parametrisations to recover unbiased cosmological constraints, using the mock data and validation procedure described here. The results of these simplified models will be briefly discussed in the next section.

4.2 Cosmological constraints from 2- and 3-point correlators

Having validated our bias model and scale cuts, we now apply them to the analysis of the real data. Our constraints on Ω_m and σ_8 for different data combinations, as well as the best-fit χ^2 and associated probability-to-exceed (PTE, i.e. the probability of obtaining a χ^2 value above that observed in the data, given the number degrees of freedom of the model), are summarised in table 2. We discuss these results in detail in the remainder of this section.

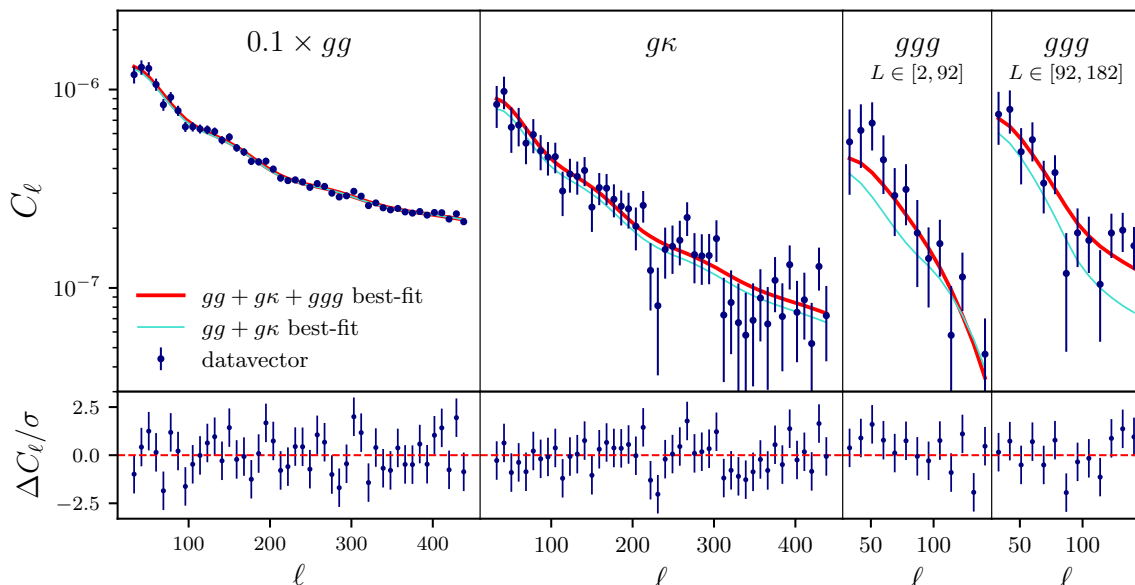


Figure 5. Power spectrum and F²B measurements, and model fit. Measurements of C_ℓ^{gg} , $C_\ell^{g\kappa}$, and $\Phi_{LL\ell}^{ggg}$ in the fourth redshift bin. Results are shown over the scales used in the analysis, with the F²B measurements covering the first two filters (the corresponding scale ranges are shown in the legend). The red line shows the best-fit model found from the combination of all data shown here, while the blue line shows the prediction found by fitting only the two-point functions. The lower panel reports the residuals with respect to the best-fit model normalized by the statistical uncertainties.

Figure 5 shows our power spectrum and F²B measurements for the fourth redshift bin, together with the best-fit theory prediction. The galaxy bispectrum is measured at high significance in the two filter scales used in the analysis (covering $2 \leq \ell < 92$ and $92 \leq \ell < 182$). Similar results are found for the other three redshift bins. As can be seen in the figure, and as quantified in table 2, we find that the model is able to fit all components of the data vector well in all cases explored here, with PTE values ranging from 0.35 to 0.78.

Figure 6 shows the two-dimensional marginalised constraints on Ω_m and σ_8 obtained from each individual redshift bin, compared with the constraints found by *Planck*. Results are shown for two-point functions only and including the F²B data. We find that the inclusion of three-point information leads to consistently smaller uncertainties by ~ 10 – 20% , in addition to a preference for a smaller value of σ_8 and larger Ω_m . Interestingly, this preference is also found in full-shape analyses of the three-dimensional clustering of galaxies when including the bispectrum [6, 77, 78]. We also find that the data in the first redshift bin favours a value of σ_8 that is significantly lower than the *Planck* constraint, by approximately 2.3σ . This is a known result, found in past analyses of this sample by [13] (S25 hereafter) and [43]. As reported in H25 [42], we find that the galaxy bispectrum measurements are also compatible with this result.

We compare these tomographic constraints on σ_8 with previous results in the top panel of figure 7. Our results are in general good agreement with those found by S25 using the same sample. Our reported uncertainties are $\sim 20\%$ larger than those of S25 for this tomographic analysis, although the latter are supplemented by external constraints on Ω_m from baryon

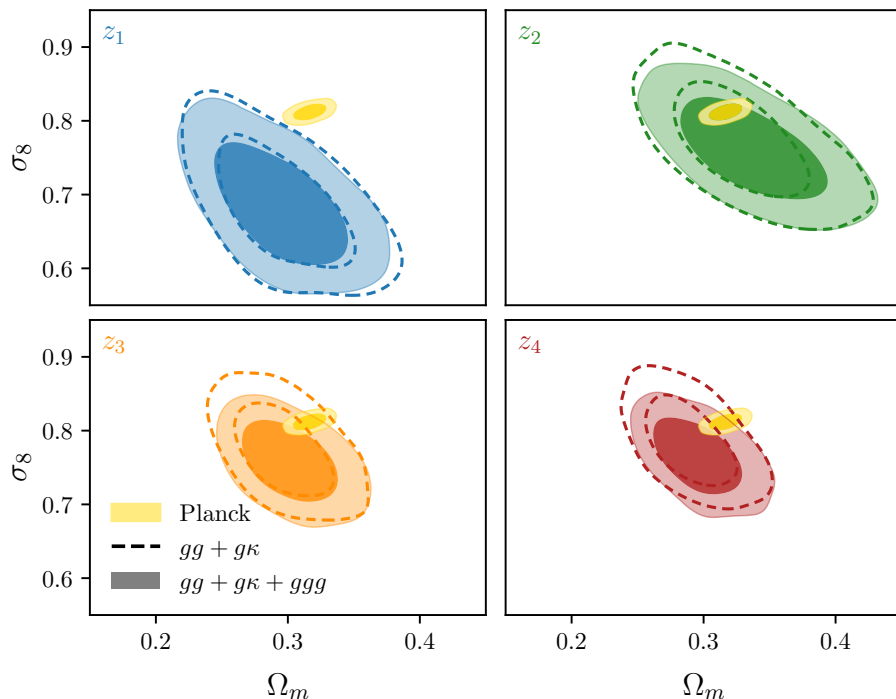


Figure 6. Cosmological constraints from each redshift bin. Two-dimensional posterior constraint on the cosmological parameters when analysing C_ℓ^{gg} and $C_\ell^{g\kappa}$ (dashed lines), and with the addition of the F²B measurements (filled contours). For reference, the Planck constraints are shown in yellow in all panels.

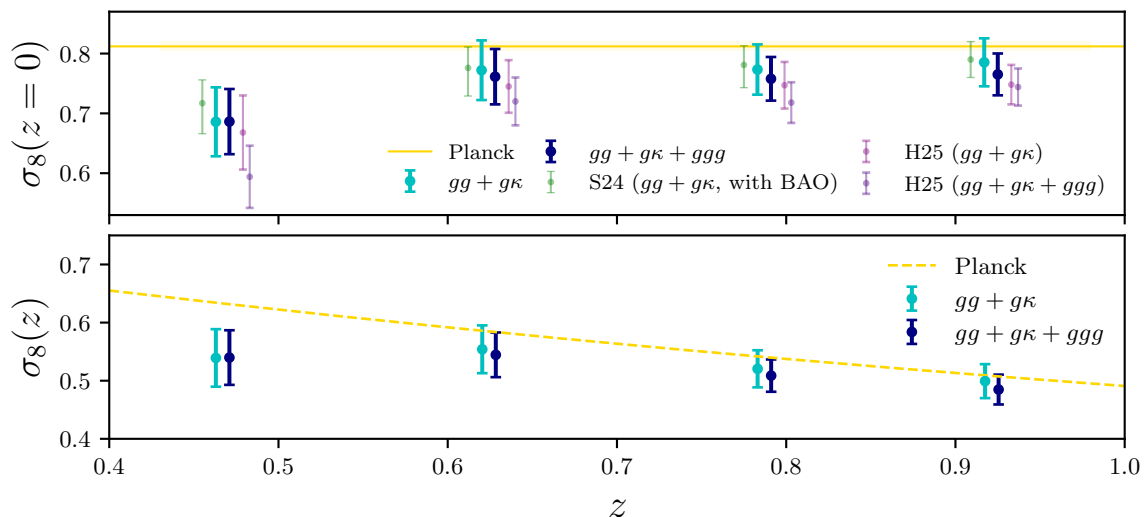


Figure 7. Tomographic constraints on σ_8 and growth. *Top:* Constraints on σ_8 found in each redshift bin using only two-point information and adding the bispectrum measurements (thick points with error bars). These results are compared with those found from the analysis of the same data in S25 [13] and H25 [42], the latter assuming coevolution relations for the bias parameters. *Bottom:* Tomographic reconstruction of the growth history in terms of $\sigma_8(z)$. The Planck best-fit prediction is shown in yellow for comparison.

acoustic oscillation data. The constraints found by H25 for two-point data are systematically lower than those found here by about $\sim 0.5\sigma$, rising to $\sim 1\sigma$ when including bispectrum data. These differences are not unexpected, however. First, H25 analysed a previous version of the LRG catalogue compared to that used here and in S25, where a number of shortcomings were addressed (see [45] for further details). Secondly, the constraints obtained by H25 from two-point data modelled the galaxy power spectrum at the tree level, equivalent to a linear bias model. Although this approach may lead to unbiased constraints using sufficiently conservative scale cuts, we saw in section 4.1 that the model is significantly less stable, and past analyses have found consistent downward shifts in amplitude parameters, such as σ_8 , when using a linear bias prescription (e.g. the “model independent” analysis of S25). Finally, when including bispectrum data, H25 made use of the coevolution relations of [79, 80] to relate the values of the higher-order bias parameters to b_1 . Any deviation from these relations would therefore result in additional bias in the inferred cosmological parameters.

Since, strictly speaking, the clustering measurements made in each redshift bin are only sensitive to the amplitude of matter fluctuations at the redshift of the galaxies that make up the sample, the quantity that is directly measured by our data is $\sigma_8(z) \equiv \sigma_8 D(z)$, where $D(z)$ is the large-scale linear growth factor.² We present the tomographic constraints on this parameter obtained from our data in the bottom panel of figure 7. Our constraints are largely compatible with the growth history predicted by *Planck* (also shown in the figure), with the exception of the mild $\sim 2\sigma$ downwards deviation observed in the first redshift bins. This is also in agreement with other constraints on the growth history from CMB lensing tomography, including S25 (using the same galaxy sample), [11] using a variety of galaxy samples, [12, 81] using the unWISE catalog, [22, 23] using high-redshift quasars from *Quaia*, [14] combining photometric LRGs and the DESI Bright Galaxy Sample and, most recently, [15] using the 2MPZ and WISE \times SuperCOSMOS galaxy samples at low redshifts. The addition of the galaxy bispectrum results in a ~ 10 – 15% reduction in the statistical uncertainties of these measurements, particularly at the higher redshifts, accompanied by the slight downwards shift in the value of $\sigma_8(z)$ discussed earlier.

Finally, we combine the measurements in different redshift bins to place joint constraints on σ_8 and Ω_m . Given the mild tension between the constraints found in the first redshift bin and the rest, we choose our fiducial constraints to correspond to the combination of the last three redshift bins, although we will also present the results of combining all redshift bins. Figure 8 shows these combined constraints, including those found using only two-point correlations, as well as those incorporating the galaxy bispectrum. Additionally, figure 9 shows the same constraints, as well as those arising from the combination of all redshift bins, in comparison with other results found in the literature. Our constraint on Ω_m is compatible with *Planck* within $\sim 1\sigma$. As shown in [15, 42, 82], this constraint is dominated by the large-scale broadband shape of the power spectrum, as constrained by C_ℓ^{gg} and $C_\ell^{g\kappa}$. Perhaps interestingly, our constraints on this parameter lie on the low side of the *Planck* measurements, as do most estimates derived from CMB lensing tomography at redshifts

²Note that, in the presence of massive neutrinos, structure growth is scale dependent at the linear level. The scale-dependent features induced by the value of M_ν assumed here ($M_\nu = 0.06$ eV) are negligible given our statistical uncertainties, however.

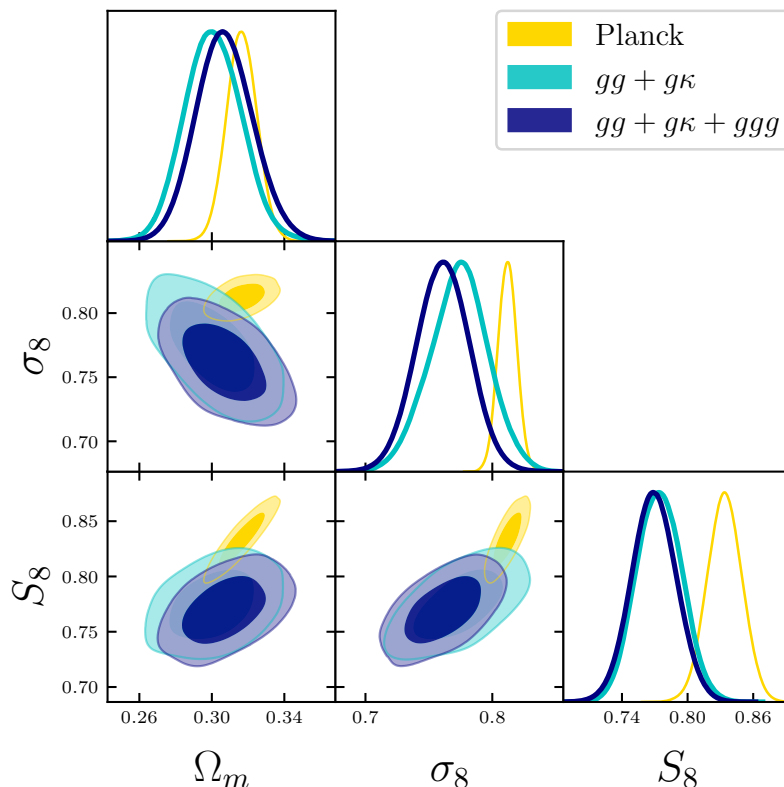


Figure 8. *Combined cosmological constraints.* Constraints on σ_8 , Ω_m , and S_8 from the combination of the three highest redshift bins analysed here. Results are shown using only two-point information (C_ℓ^{gg} and $C_\ell^{g\kappa}$, *turquoise*), and with the addition of the bispectrum measurements ($\Phi_{LL\ell}^{ggg}$, *dark blue*).

$z \lesssim 1$, as well as the latest BAO constraints on Λ CDM from DESI [83]. Nevertheless, the measurement uncertainties are too large to draw any definitive conclusion regarding a potential parameter tension.

In turn, our fiducial constraints on σ_8 are lower than *Planck* at the 2.4σ level when including power spectrum and bispectrum measurements. Compared to the 1.6σ deviation obtained using only two-point functions, this result is driven by the $\sim 15\%$ smaller error bars combined with the $\sim 0.5\sigma$ downward shift in σ_8 we discussed above. Combining the measurements from all redshift bins, this tension in the value of σ_8 rises to the 3σ level (2.5σ from two-point functions only), driven by the significantly lower amplitude of matter fluctuations favoured by the first redshift bin.

We can also derive constraints on the so-called “clumping parameter” $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$, which can be accurately constrained by cosmic shear analyses. Combining different redshift bins, we obtain:

$$S_8 = 0.769 \pm 0.020 (z_2 + z_3 + z_4), \quad S_8 = 0.756 \pm 0.018 (\text{all bins}). \quad (4.2)$$

Compared to the *Planck* constraints on this parameter ($S_8 = 0.832 \pm 0.013$), these measurements lie 2.6 and 3.4σ lower, respectively. The tension in the latter case is driven by the inclusion of the first redshift bin, where the constraints on both Ω_m and σ_8 lie below the *Planck* values (see top left panel of figure 6). This level of tension is marginally higher

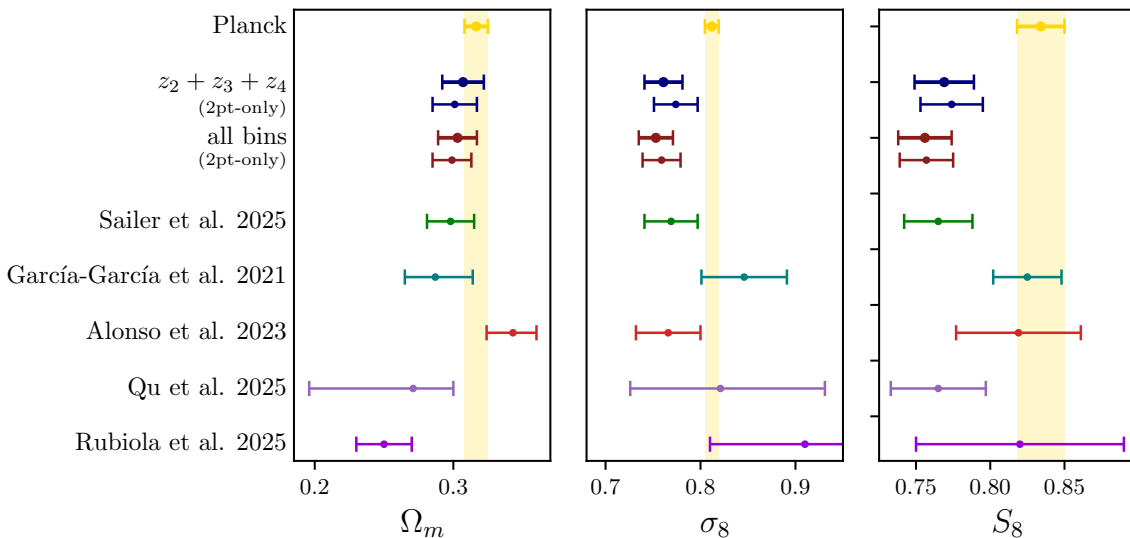


Figure 9. *Comparison with past constraints.* Comparison of our constraints on $\{\Omega_m, \sigma_8, S_8\}$ with past constraints derived from lensing tomography [11, 13, 15, 82, 84] as well as the *Planck* results.

than that found in [13] using lensing from *Planck* and ACT (2.3σ), driven primarily by the shift induced by the addition of bispectrum data, as well as the different modelling choices. Figure 9 shows our constraints on Ω_m and σ_8 in the context of other CMB lensing tomography analyses, including S25 [13], the analysis of [11] combining multiple galaxy samples, the constraints of [84] using an alternative sample from the Legacy Survey, the recent analysis of [15] using low-redshift data, and the high-redshift constraints of [82] using quasars from *Quaia*. The constraints found here are in broad agreement with these works, although it is interesting to note that the analyses showing the strongest tension in σ_8/S_8 with the CMB are those using LRG samples. Additionally, with the exception of [82], most lensing tomography analyses, in the absence of external BAO data, seem to favour a marginally low value of Ω_m . Given the current uncertainties, this could be simply a mild statistical anomaly, although misspecification in the non-linear bias model, or the presence of systematic contamination in the observed fluctuations in galaxy number counts, could also contribute to a low S_8 or Ω_m .

To further explore the limits of the bias model used here, and the ability of the bispectrum measurements to improve the final cosmological constraints, we have explored the possibility of using simpler bias models to describe the data. In particular, we considered two cases: One in which the counter-term parameters $\{c_m, b_{k^2}\}$ and the third-order bias parameter b_{3nl} are set to zero, and one in which the higher-order bias parameters $\{b_2, b_s, b_{3nl}\}$ are set to zero. In the first case, the power spectra and bispectra are both described by the same set of bias parameters, thus maximising the information content in the bispectrum. The details of this study are discussed in appendix A.2, where we show that these simplified models are able to describe the data well over the same range of scales as our fiducial bias model. When applied to the real data, we (unsurprisingly) recover tighter constraints on both Ω_m and σ_8 , with a small shift towards the values preferred by *Planck*. Interestingly, as shown in table 2, both models (indicated by the subset of parameters that we vary) are able to fit the data well,

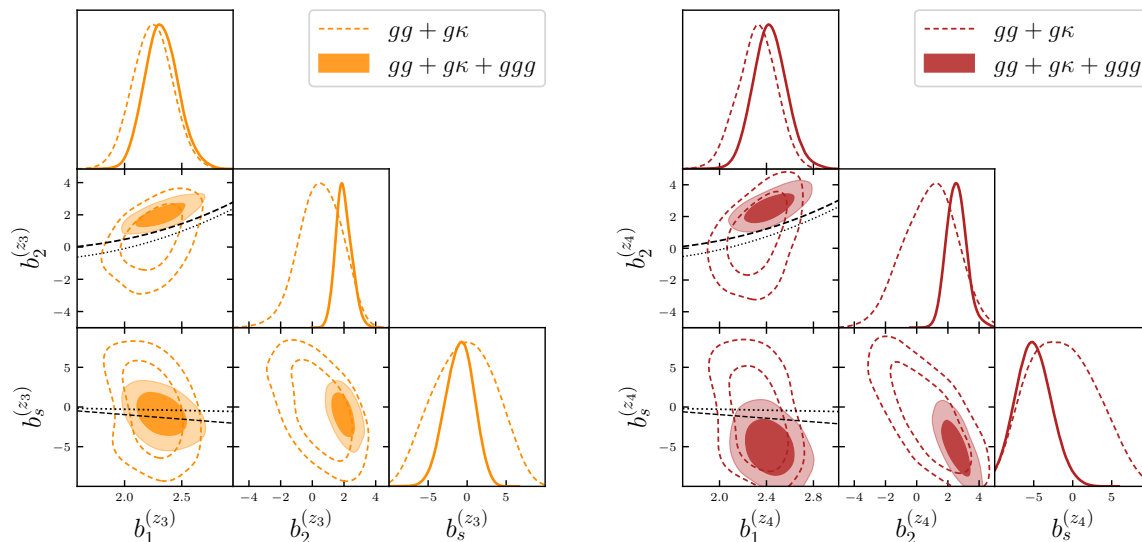


Figure 10. *Constraints on bias parameters.* Impact of additional three-point information on the posterior of the linear and higher-order bias parameters $\{b_1, b_2, b_s\}$ that contribute to our model of the bispectrum (eq. (3.13)). Results are shown for the two highest redshift bins, with similar results found for the first two bins. The inclusion of three-point information leads to a factor $\mathcal{O}(5\text{--}10)$ improvement in the FoM for this set of bias parameters. The coevolution relations obtained from [79, 80] for haloes are shown as dotted lines, while the relations found by [85] for galaxies are shown as dashed lines.

with a very small penalty in the best-fit χ^2 with respect to the fiducial model, commensurate with the reduced number of degrees of freedom. This suggests the additional complexity of the fiducial model is not required by the data. Nevertheless, a more detailed validation of these simpler models, and their ability to describe the clustering of generic galaxy samples, would be required before they can be used reliably in cosmological analyses.

4.3 Galaxy bias relations from projected clustering

Galaxy bias is a fundamental element of perturbation theory models for non-linear clustering. In this first-principles approach, the bias coefficients have to be treated as nuisance parameters in cosmological inference, introducing degeneracies that inevitably limit the precision of parameter estimation. Moreover, non-linear bias parameters notably contribute to the power spectrum in a largely degenerate way, making it challenging to independently constrain them using power spectrum alone in 3D analyses. This problem becomes more pronounced for projected clustering, where linear features such as the BAO wiggles are highly suppressed. In this context, the inclusion of higher-order correlation functions helps mitigate this limitation: The need to simultaneously describe two- and three-point functions helps break the degeneracies between different bias parameters, allowing us to measure them at higher precision. A better constraint on the higher-order bias parameters may then allow us to use the galaxy-galaxy and galaxy- κ power spectra on mildly non-linear scales to constrain cosmology.

Figure 10 shows the constraints on the bias parameters using only power spectra or power spectra and bispectrum measurements. Results are shown for the last two redshift bins, where we are able to use a larger number of angular modes at a fixed k_{\max}^B and

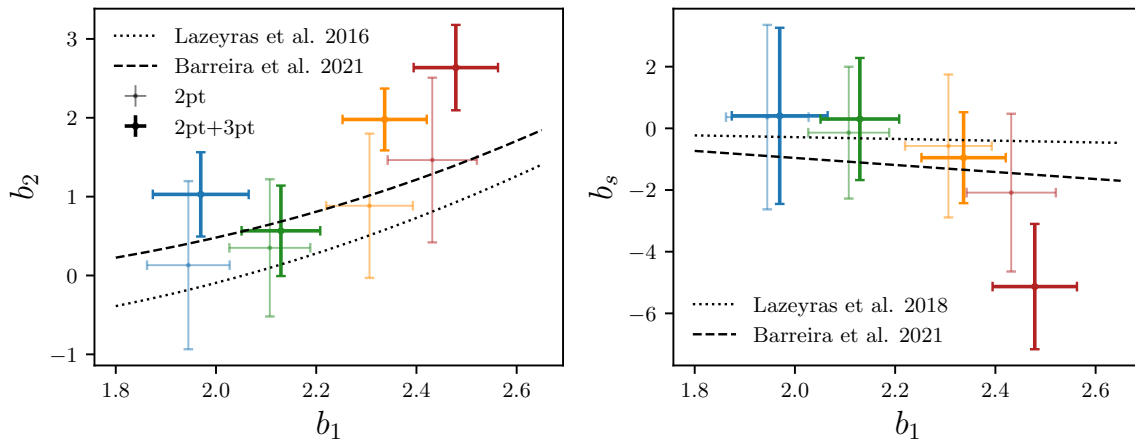


Figure 11. *Bias relations.* Relation between the quadratic (b_2 , left) and tidal (b_s , right) bias parameters and the linear bias b_1 found in each redshift bin. Results are shown using only two-point information (*thinner, shaded markers*), and with the inclusion of bispectrum data (*thicker markers*). These constraints are compared against the coevolution relations for haloes ([79, 80], *dotted lines*), and galaxies ([85], *dashed lines*).

where, in consequence, this improvement is more evident. Quantitatively, the inclusion of bispectrum information, even on the conservatively large scales used here, leads to a remarkable $\mathcal{O}(10)$ improvement in the FoM for the three bias parameters $\{b_1, b_2, b_s\}$, with b_2 and b_s experiencing the strongest reduction in their statistical uncertainties. Importantly, the constraints obtained from the full data vector are compatible with those obtained from power spectra alone, with the bispectrum favouring the region of parameter space with larger values of b_1 and b_2 (and marginally lower b_s). This trend leads to the mild downwards shift in σ_8 observed in the previous section when including bispectrum data (as also observed in [78]). Also, and reassuringly, the best-fit bias parameters $\{b_1, b_2, b_s\}$ found using power spectrum data alone provide a reasonable prediction for the bispectrum measurements (blue lines in figure 5), demonstrating the internal self-consistency of our data and model. Having improved constraints allows us to examine the relationships between the different bias parameters. In particular, measurements of the clustering of galaxies and haloes in N -body simulations have found that the higher-order bias parameters exhibit a correlation with the linear bias b_1 . These often called ‘‘coevolution relations’’ are not completely universal, differing slightly between haloes and galaxies, and showing a non-negligible scatter. Nevertheless, it is interesting to confirm whether the clustering of galaxies observed in real data also obeys these relations. Figure 10 shows the relations found by [79] and [80] for haloes (dotted lines), and by [85] for galaxies (dashed lines). The preference of the bispectrum data for a larger b_2 pushes our constraints marginally away from the coevolution relations. This is shown in more detail in figure 11, which collects the constraints found in all redshift bins when analysed jointly. With the inclusion of bispectrum data, our constraints show a mild preference (at the $\sim 2\sigma$ level) for upwards (downwards) deviations in b_2 (b_s) for larger values of b_1 . These deviations are not entirely unexpected, given the relatively large scatter found in the case of galaxies [85–87]. Note also that our constraints use data from a single galaxy type (LRGs), spanning a relatively narrow range of b_1 values.

5 Conclusions

We have carried out the first comprehensive cosmological analysis of lensing tomography incorporating information from the angular galaxy bispectrum, complementing the first tentative analysis of [42]. A key aspect of such an analysis is the high redundancy of the full data vector, forcing our model of the galaxy-matter connection to simultaneously describe the two- and three-point clustering of galaxies as well as their cross-correlation with the matter overdensities. This redundancy improves both the precision and robustness of the resulting cosmological constraints.

We have used a perturbative bias expansion within the Effective Field Theory of LSS framework, modelling the galaxy-galaxy and galaxy-convergence power spectra at one-loop order, and the galaxy bispectrum at tree level. We have validated this model against N -body-based mock observations, and determined the scales over which it can safely recover cosmological constraints. Compared to studies of the three-dimensional clustering of galaxies, this task is made simpler in the case of projected statistics due to the negligible impact of non-linear redshift-space distortions. We have shown that the model is able to recover unbiased constraints down to scales where cosmological information is in fact saturated for the LRG sample studied here.

Applying this model to the analysis of the four different redshift bins considered here, we find constraints on σ_8 and Ω_m that are consistent with past works analysing this sample, and have shown that the inclusion of the galaxy bispectrum leads to a ~ 10 – 15% improvement on the cosmological constraints. The constraints on Ω_m and σ_8 found from the three highest-redshift bins are compatible with each other and in reasonable agreement with *Planck*, as is the growth history reconstructed from these data (see figure 7). In turn, the first redshift bin shows a preference for lower σ_8 , and is in tension with *Planck* at the $\sim 2.5\sigma$ level.

We have then derived cosmological constraints from the joint analysis of the last three redshift bins. The measured value of σ_8 (S_8) is moderately lower than the *Planck* value at the 2.4σ (2.6σ) level. This is due to the mild but coherent preference for a lower σ_8 displayed by the two-point clustering of galaxies in all redshift bins, combined with the additional $\sim 0.5\sigma$ downwards shift observed when including the galaxy bispectrum. Interestingly, a similar shift has been observed in the three-dimensional clustering of galaxies when combining power spectrum and bispectrum data [78]. This tension is then exacerbated, above the 3σ level, when combined with the first redshift bin. In agreement with the results obtained in our per-bin analysis, the inclusion of the projected bispectrum leads to a $\sim 10\%$ improvement in the statistical uncertainties on cosmological parameters.

Including the bispectrum has a much stronger impact on the higher-order bias parameters of our model, b_2 and b_s , with a reduction in their statistical uncertainties of more than a factor 2 in some cases. This allows us to study the relations between the different bias parameters with higher precision, as well as their agreement with coevolution relations found in simulations. The constraints obtained after including bispectrum information are in good agreement with those from power spectra alone, but show a consistent preference for higher b_1 and b_2 , and lower b_s . This manifests itself as a mild departure from the coevolution relations of [85] for high values of b_1 . This is not entirely surprising, however, given the large scatter expected around these relations.

Adding bispectrum data improves the robustness and precision of the resulting cosmological constraints. The improvement in the latter case is relatively disappointing, however ($\sim 10\text{--}15\%$), and it is interesting to reflect on the potential reasons for this. First, although the bispectrum is able to significantly improve the constraints on b_2 and b_s when modelled at the tree level, the counter-term parameters $\{b_{k^2}, c_m\}$ and the third-order parameter b_{3nl} are only constrained by the power spectrum measurements. This leaves significant freedom to describe the featureless clustering of galaxies on mildly non-linear scales. In appendix A.2 we have explored the possibility of using a simpler bias model, setting $\{b_{k^2}, c_m, b_{3nl}\}$ to zero, and thus allowing us to maximally exploit the information contained in the bispectrum measurements. Interestingly, we have shown that this simplified model is still able to recover unbiased constraints in mock data and, when applied to the data, the aforementioned tensions are somewhat relieved. Nevertheless, the relative improvement in the final constraints from including the bispectrum is still limited to $\sim 10\%$, suggesting that the limiting factor is in fact not the additional freedom allowed by the counterterm parameters and b_{3nl} .

Another potential cause for the mild improvement from the bispectrum is the large uncertainties in $C_\ell^{g\kappa}$ and $\Phi_{LL\ell}^{ggg}$ relative to the galaxy power spectrum. Both of these measurements are crucial to break degeneracies in the bias model and constrain cosmology, and the final constraints may be limited by these uncertainties, rather than freedom in the model itself. Future work could improve both of these aspects of our analysis using existing data. First, replacing the CMB lensing convergence with cosmic shear data from existing imaging surveys [88, 89] could improve the sensitivity of the galaxy-lensing power spectrum. It could also allow us to exploit the galaxy-galaxy-lensing bispectrum, which we have discarded in this analysis due to the low signal-to-noise ratio on the large scales over which it can be modelled accurately, as shown in [42]. Secondly, our measurements of the angular bispectrum have been based on the F²B estimator presented in [41], which only recovers triangle configurations that are close to isosceles. As discussed in section 2.2 and [41], the F²B estimator can be easily generalised to recover the remaining non-zero bispectrum configurations, thus increasing the total three-point statistical power. Finally, our use of the bispectrum is limited by the range of scales over which our tree-level model is sufficiently accurate. In the future, this could be significantly improved by resorting to more sophisticated approaches to galaxy bias, such as hybrid effective field theory (HEFT, [90–92]), which could remain accurate on significantly smaller scales without a substantial increase in the number of bias parameters required.

This work demonstrates that a joint analysis of the projected power spectrum and bispectrum of galaxies and weak lensing can be efficiently and reliably carried out with existing data. The deployment of these methods could also enhance the science returns of future imaging surveys, such as the Rubin Observatory [26, 27], maximising the information extracted from the projected clustering of galaxies.

Acknowledgments

We thank Carlos García-García, Andrina Nicola, Sergi Novell-Massot, Noah Sailer, Emiliano Sefusatti and Anže Slosar for useful comments and discussions. LH is supported by a Hintze studentship, which is funded through the Hintze Family Charitable Foundation. MZ and

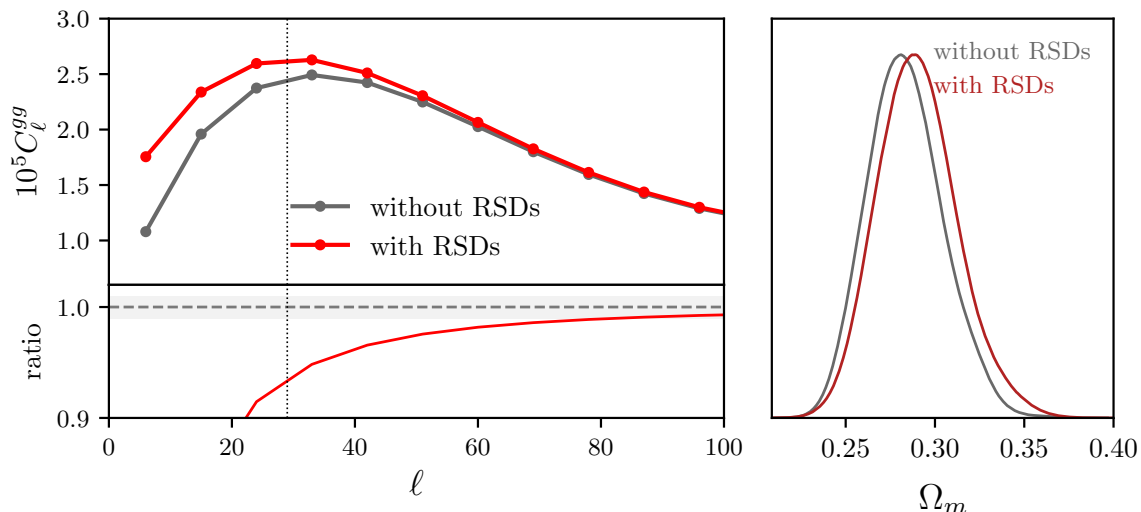


Figure 12. *The impact of RSDs.* *Left:* Contribution of RSDs to the galaxy power spectrum on large scales. The vertical dotted line represents the large-scale cut that we adopt in our fiducial analysis. *Right:* Systematic shift on the inferred Ω_m parameter for the LRGs z_3 bin only when ignoring the RSD contribution. Using our fiducial large-scale cut, the shift is relatively minor ($\sim 0.3\sigma$).

DA acknowledge support from the Beecroft trust. We made extensive use of computational resources at the University of Oxford Department of Physics, funded by the John Fell Oxford University Press Research Fund.

A Additional tests

A.1 Treatment of redshift-space distortions

Redshift-space distortions are the most relevant secondary effect modifying the galaxy angular power spectrum on large scales, particularly for relatively narrow redshift bins [57, 93]. As described in section 3.2.1, we include RSDs at the linear level in the galaxy auto-spectrum, including only the correlations between the RSD term and the linear density term proportional to b_1 and the RSD auto-correlation. Given the results of our simulation-based validation, this is sufficient to obtain unbiased cosmological constraints from our data. Nevertheless, to gain a better intuitive understanding of the impact of RSDs on our analysis, the left panel of figure 12 shows the large-scale power spectrum with and without the RSD contributions. Within the scale cuts used here, RSDs lead to deviations in the power spectrum amplitude of up to $\sim 10\%$, peaking on the largest scales used and becoming negligible at high ℓ .

Repeating our analysis omitting these contributions, the right panel of figure 12 shows that the largest effect is a relatively mild ($\sim 0.3\sigma$) downwards shift in the preferred value of Ω_m . Thus, although relevant for this analysis, a precise modelling of RSDs (e.g. including the impact of non-linear “fingers of God”) is not required.

A.2 Simpler bias parametrisations

The bias model used here (see section 3.2.3), in which power spectra are modelled at the 1-loop level including EFT counterterms, and the bispectrum is modelled at the tree level,

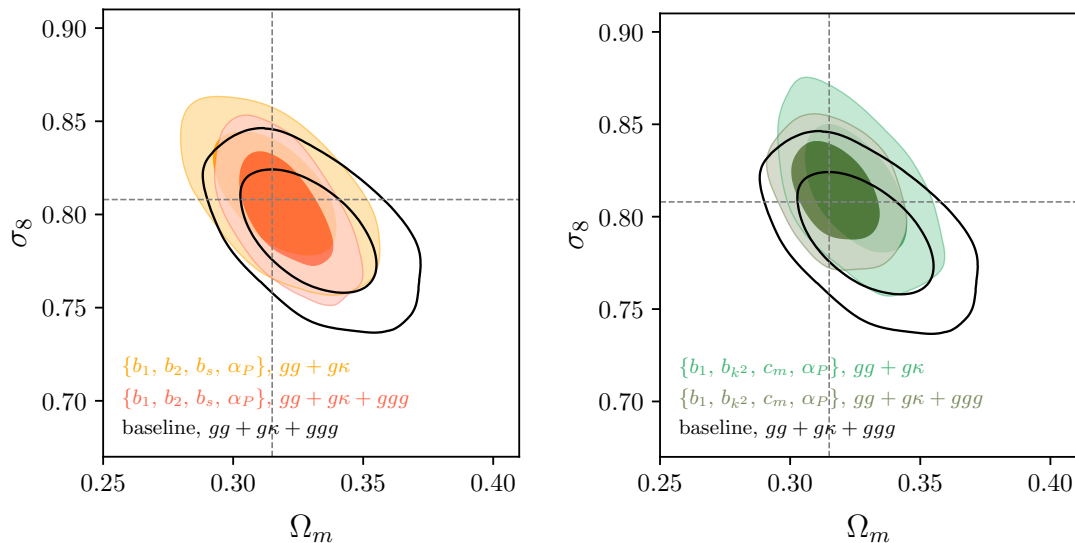


Figure 13. *Simpler bias models on simulations.* Constraints on Ω_m and σ_8 obtained from the ABACUSUMMIT simulation, combining all redshift bins. Our baseline constraints are shown as black lines. The brighter and darker contours show the constraints found using only two-point functions and combining them with the bispectrum, respectively. The *left panel* shows the constraints found using a simpler bias model in which the scale-dependent counterterms $\{b_{k^2}, c_m\}$, and the third-order parameter b_{3nl} are set to zero. The right panel shows the constraints found setting b_2 , b_s , and b_{3nl} to zero instead, in a model inspired by the phenomenological parametrisation of [20]. Both restricted parametrizations are able to find unbiased constraints on the cosmological parameters.

involves free parameters, such as c_m , b_{k^2} , and b_{3nl} , which can only be constrained by a single type of correlation (e.g. two-point functions, but not the bispectrum). Although this level of freedom is theoretically necessary and physically motivated, it is interesting to explore whether simpler versions of this model are still able to provide unbiased cosmological constraints. Furthermore, recent works [20] have shown that, in the case of real-space galaxy clustering, a simplified phenomenological parametrization of galaxy bias is still able to provide an adequate and unbiased description of the data. In this section we explore how incorporating bispectrum information impacts parameter constraints when working within two restricted bias parametrizations.

Specifically, we test two complementary simplified models. First, we explore a parametrization that only retains the parameters that enter in eq. (3.13), i.e. varying $\{b_1, b_2, b_s, \alpha_P\}$, while fixing the counterterm values, and third order Eulerian bias, to zero. Then, we also tested a theoretical model that, in the spirit of [20], effectively only parametrizes the deviations from a scale-independent linear bias through counterterm-like coefficients. In our setup, we achieve this by only varying the subset of nuisance parameters $\{b_1, b_{k^2}, c_m, \alpha_P\}$, fixing to zero the remaining bias parameters.³

We begin by testing its performance on simulated data. As shown in figure 13, we find that both simplified parametrizations are able to recover the fiducial values of the cosmological pa-

³Note that the model in [20] also allows for a stochastic noise component in the galaxy-matter power spectrum, making it more general than the model studied here.

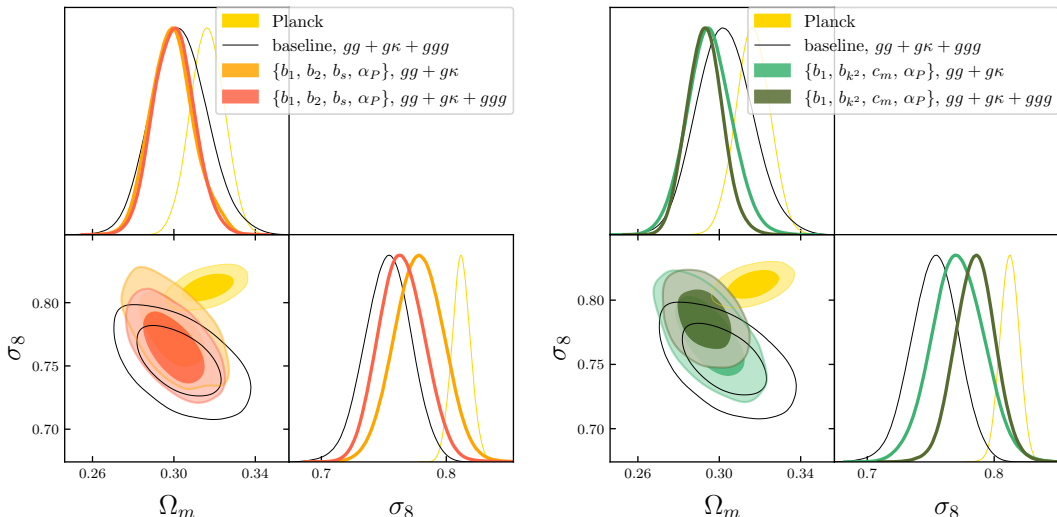


Figure 14. *Simpler bias models on real data.* As figure 13, now showing the constraints found when analysing the real data. The *Planck* constraints are also shown for comparison.

parameters when applied to the combination of all redshift bins, both using only two-point correlators and including the bispectrum. This is likely due to the relatively featureless nature of the projected galaxy statistics, making it possible to describe them with a relatively small number of smooth “basis” functions. Notably, in this simulated data, the addition of *ggg* information is perfectly compatible with the 2-point posterior, tightening the constraints around the (Ω_m, σ_8) fiducial point, improving their statistical uncertainties by (26%, 8%) using the $\{b_1, b_2, b_s, \alpha_P\}$ model (left panel) and (17%, 24%) using the $\{b_1, b_{k^2}, c_m, \alpha_P\}$ model (right panel).

Then, having validated both on simulations, we apply them to the analysis of the real clustering data. Figure 14 shows the posterior distribution of the cosmological parameters, as well as the constraints found using our baseline model for reference. We find that, using the $\{b_1, b_2, b_s, \alpha_P\}$, the uncertainty on Ω_m shrinks by $\sim 25\%$ with respect to our baseline model, while the σ_8 remains largely unchanged. The value also favours a slightly lower value of Ω_m and a larger value of σ_8 , reducing the tension with *Planck* found in our baseline analysis. A similar shift is found in the case of the $\{b_1, b_{k^2}, c_m, \alpha_P\}$ model, again relieving this tension. The inclusion of the bispectrum in this model also leads to a significant $\sim 20\%$ reduction in the σ_8 uncertainty.

It is worth stressing the fact that these two simpler bias models, defined as subsets of the full EFT bias model presented, are not, strictly speaking, well motivated or complete from a physical point of view. Thus, their ability to describe our data, as tested on simulations, might be purely accidental and only applicable to the specific galaxy sample we have simulated, which would lead to artificially reduced cosmological constraints. Therefore, a more comprehensive validation of the models, considering a large suite of non-linear galaxy clustering scenarios, would be necessary before they can be safely applied to cosmological analyses.

A.3 Emulators accuracy

The approach of emulating, rather than computing *ab initio* at each likelihood step, the theory templates enables fast and feasible MCMC sampling as detailed in section 3.3. While

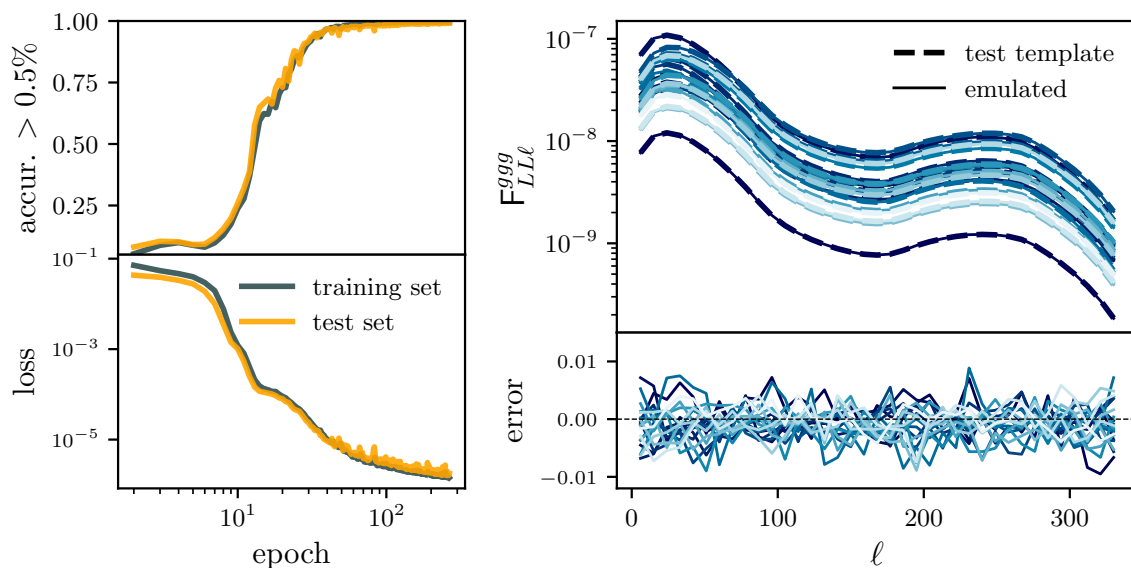


Figure 15. *Emulating F^2B templates.* Training history (*left*) of the emulator for the $F_{LL\ell}^{ggg}$ (first filter) template, and performance (*right*) on some elements of the test set. All the emulators for the other templates present an analogue training history and a similar performance.

essential, this additional step may introduce systematic errors; therefore, the accuracy of the emulators must be carefully validated.

To this purpose, we focus on the emulator of the $F_{LL\ell}^{ggg}$ template as a representative example, as it exhibits consistent behaviour with the others. Specifically, in figure 15 we show the evolution of the loss function (mean squared errors) and of our accuracy metric throughout the training epochs. We find that in a few hundred epochs the loss function improves by almost 5 orders of magnitude, and virtually 100% of the output layer neurons differ less than 0.5% from the ground truth (which prompts us to stop the training). We repeat this exercise for both the training set and the test set (which is never seen by the Neural Network while training), finding similar results and no indication of overfitting.

Finally, in the right panel of figure 15, we show the overall accuracy of this emulator, in terms of percentage differences between a subsample of the $F_{LL\ell}^{ggg}$ template values used in the test set and their corresponding emulator predictions. We find that the emulator reproduces the test set at better than 1% over the whole considered ℓ range, with no significant scale dependence. We find similar performance for the emulators of all of the templates of section 3.3, which is accurate enough given the statistical error of the dataset used in this work.

References

- [1] N. Kaiser, *On the Spatial correlations of Abell clusters*, *Astrophys. J. Lett.* **284** (1984) L9 [INSPIRE].
- [2] G. Efstathiou, W.J. Sutherland and S.J. Maddox, *The cosmological constant and cold dark matter*, *Nature* **348** (1990) 705 [INSPIRE].
- [3] J.A. Peacock, *The power spectrum of galaxy clustering*, *Mon. Not. Roy. Astron. Soc.* **253** (1991) 1P.

- [4] J.A. Peacock and D. Nicholson, *The large-scale clustering of radio galaxies*, *Mon. Not. Roy. Astron. Soc.* **253** (1991) 307.
- [5] C.M. Baugh and G. Efstathiou, *The three-dimensional power spectrum measured from the APM Galaxy Survey — I. Use of the angular correlation function*, *Mon. Not. Roy. Astron. Soc.* **265** (1993) 145.
- [6] O.H.E. Philcox and M.M. Ivanov, *BOSS DR12 full-shape cosmology: Λ CDM constraints from the large-scale galaxy power spectrum and bispectrum monopole*, *Phys. Rev. D* **105** (2022) 043517 [[arXiv:2112.04515](#)] [[INSPIRE](#)].
- [7] M.M. Ivanov et al., *Full-shape analysis with simulation-based priors: Constraints on single field inflation from BOSS*, *Phys. Rev. D* **110** (2024) 063538 [[arXiv:2402.13310](#)] [[INSPIRE](#)].
- [8] DESI collaboration, *DESI 2024 V: Full-Shape galaxy clustering from galaxies and quasars*, *JCAP* **09** (2025) 008 [Erratum *ibid.* **02** (2026) E02] [[arXiv:2411.12021](#)] [[INSPIRE](#)].
- [9] S. Novell-Masot et al., *Full-Shape analysis of the power spectrum and bispectrum of DESI DR1 LRG and QSO samples*, *JCAP* **06** (2025) 005 [[arXiv:2503.09714](#)] [[INSPIRE](#)].
- [10] A. Nicola et al., *Tomographic galaxy clustering with the Subaru Hyper Suprime-Cam first year public data release*, *JCAP* **03** (2020) 044 [[arXiv:1912.08209](#)] [[INSPIRE](#)].
- [11] C. García-García et al., *The growth of density perturbations in the last ~ 10 billion years from tomographic large-scale structure data*, *JCAP* **10** (2021) 030 [[arXiv:2105.12108](#)] [[INSPIRE](#)].
- [12] G.S. Farren et al., *The Atacama Cosmology Telescope: Cosmology from Cross-correlations of unWISE Galaxies and ACT DR6 CMB Lensing*, *Astrophys. J.* **966** (2024) 157 [[arXiv:2309.05659](#)] [[INSPIRE](#)].
- [13] N. Sailer et al., *Cosmological constraints from the cross-correlation of DESI Luminous Red Galaxies with CMB lensing from Planck PR4 and ACT DR6*, *JCAP* **06** (2025) 008 [[arXiv:2407.04607](#)] [[INSPIRE](#)].
- [14] N. Sailer et al., *Evolution of structure growth during dark energy domination: Insights from the cross-correlation of DESI galaxies with CMB lensing and galaxy magnification*, *Phys. Rev. D* **111** (2025) 103540 [[arXiv:2503.24385](#)] [[INSPIRE](#)].
- [15] A. Rubiola, M. Zennaro, C. García-García and D. Alonso, *Low-redshift constraints on structure growth from CMB lensing tomography*, [arXiv:2510.09563](#) [[INSPIRE](#)].
- [16] Y.-K. Chiang, B. Ménard and D. Schiminovich, *Broadband Intensity Tomography: Spectral Tagging of the Cosmic UV Background*, *Astrophys. J.* **877** (2019) 150 [[arXiv:1810.00885](#)] [[INSPIRE](#)].
- [17] N. Koukoufilippas, D. Alonso, M. Bilicki and J.A. Peacock, *Tomographic measurement of the intergalactic gas pressure through galaxy-tSZ cross-correlations*, *Mon. Not. Roy. Astron. Soc.* **491** (2020) 5464 [[arXiv:1909.09102](#)] [[INSPIRE](#)].
- [18] Y.-K. Chiang, R. Makiya, B. Ménard and E. Komatsu, *The Cosmic Thermal History Probed by Sunyaev-Zeldovich Effect Tomography*, *Astrophys. J.* **902** (2020) 56 [[arXiv:2006.14650](#)] [[INSPIRE](#)].
- [19] Y.-K. Chiang, R. Makiya and B. Ménard, *Cosmic Infrared Background Tomography and a Census of Cosmic Dust and Star Formation*, *Astrophys. J.* **992** (2025) 65 [[arXiv:2504.05384](#)] [[INSPIRE](#)].
- [20] S. Maleubre et al., *The impact of galaxy bias on cross-correlation tomography*, *Mon. Not. Roy. Astron. Soc.* **545** (2025) staf2125 [[arXiv:2508.05319](#)] [[INSPIRE](#)].

- [21] H. Miyatake et al., *First Identification of a CMB Lensing Signal Produced by 1.5 Million Galaxies at $z \sim 4$: Constraints on Matter Density Fluctuations at High Redshift*, *Phys. Rev. Lett.* **129** (2022) 061301 [[arXiv:2103.15862](#)] [[INSPIRE](#)].
- [22] G. Piccirilli et al., *Growth history and quasar bias evolution at $z < 3$ from Quiaia*, *JCAP* **06** (2024) 012 [[arXiv:2402.05761](#)] [[INSPIRE](#)].
- [23] C. Embil Villagra et al., *The Atacama Cosmology Telescope: high-redshift measurement of structure growth from the cross-correlation of Quiaia quasars and CMB lensing from ACT DR6 and Planck PR4*, *JCAP* **12** (2025) 033 [[arXiv:2507.08798](#)] [[INSPIRE](#)].
- [24] A. Albrecht et al., *Report of the Dark Energy Task Force*, [astro-ph/0609591](#) [[INSPIRE](#)].
- [25] R. Laureijs et al., *Euclid Definition Study Report*, [arXiv:1110.3193](#) [[INSPIRE](#)].
- [26] Ž. Ivezić et al., *LSST: from Science Drivers to Reference Design and Anticipated Data Products*, *Astrophys. J.* **873** (2019) 111 [[arXiv:0805.2366](#)] [[INSPIRE](#)].
- [27] LSST DARK ENERGY SCIENCE collaboration, *The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document*, [arXiv:1809.01669](#) [[DOI:10.2172/1471560](#)] [[INSPIRE](#)].
- [28] D. Spergel et al., *Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*, [arXiv:1503.03757](#) [[INSPIRE](#)].
- [29] H. Gil-Marín et al., *The power spectrum and bispectrum of SDSS DR11 BOSS galaxies — I. Bias and gravity*, *Mon. Not. Roy. Astron. Soc.* **451** (2015) 539 [[arXiv:1407.5668](#)] [[INSPIRE](#)].
- [30] H. Gil-Marín et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies*, *Mon. Not. Roy. Astron. Soc.* **465** (2017) 1757 [[arXiv:1606.00439](#)] [[INSPIRE](#)].
- [31] N.S. Sugiyama, S. Saito, F. Beutler and H.-J. Seo, *Towards a self-consistent analysis of the anisotropic galaxy two- and three-point correlation functions on large scales: application to mock galaxy catalogues*, *Mon. Not. Roy. Astron. Soc.* **501** (2021) 2862 [[arXiv:2010.06179](#)] [[INSPIRE](#)].
- [32] O.H.E. Philcox et al., *Cosmology with the redshift-space galaxy bispectrum monopole at one-loop order*, *Phys. Rev. D* **106** (2022) 043530 [[arXiv:2206.02800](#)] [[INSPIRE](#)].
- [33] M. Bucher, B. Racine and B. van Tent, *The binned bispectrum estimator: template-based and non-parametric CMB non-Gaussianity searches*, *JCAP* **05** (2016) 055 [[arXiv:1509.08107](#)] [[INSPIRE](#)].
- [34] PLANCK collaboration, *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9 [[arXiv:1905.05697](#)] [[INSPIRE](#)].
- [35] DES collaboration, *Dark Energy Survey Year 3 results: Cosmology with moments of weak lensing mass maps*, *Phys. Rev. D* **106** (2022) 083509 [[arXiv:2110.10141](#)] [[INSPIRE](#)].
- [36] S. Heydenreich, L. Linke, P. Burger and P. Schneider, *A roadmap to cosmological parameter analysis with third-order shear statistics — I. Modelling and validation*, *Astron. Astrophys.* **672** (2023) A44 [[arXiv:2208.11686](#)] [[INSPIRE](#)].
- [37] G. Jung, F. Oppizzi, A. Ravenni and M. Liguori, *The integrated angular bispectrum*, *JCAP* **06** (2020) 035 [[arXiv:2004.03574](#)] [[INSPIRE](#)].
- [38] P. Chakraborty, S.-F. Chen and C. Dvorkin, *Skewing the CMB \times LSS: a fast method for bispectrum analysis*, *JCAP* **07** (2022) 038 [[arXiv:2202.11724](#)] [[INSPIRE](#)].
- [39] O.H.E. Philcox, *Optimal estimation of the binned mask-free power spectrum, bispectrum, and trispectrum on the full sky: Scalar edition*, *Phys. Rev. D* **107** (2023) 123516 [[arXiv:2303.08828](#)] [[INSPIRE](#)].

- [40] G.S. Farren et al., *Detection of the CMB lensing — galaxy bispectrum*, [arXiv:2311.04213](#) [[INSPIRE](#)].
- [41] L. Harscouet et al., *Fast Projected Bispectra: the filter-square approach*, *Open J. Astrophys.* **8** (2025) 6 [[arXiv:2409.07980](#)] [[INSPIRE](#)].
- [42] L. Harscouet, D. Alonso, A. Nicola and A. Slosar, *Constraints from CMB lensing tomography with projected bispectra*, [arXiv:2507.07968](#) [[DOI:10.33232/001c.155341](#)] [[INSPIRE](#)].
- [43] M. White et al., *Cosmological constraints from the tomographic cross-correlation of DESI Luminous Red Galaxies and Planck CMB lensing*, *JCAP* **02** (2022) 007 [[arXiv:2111.09898](#)] [[INSPIRE](#)].
- [44] H. Xu et al., *DESI Legacy Imaging Surveys Data Release 9: Cosmological constraints from galaxy clustering and weak lensing using the minimal bias model*, *Sci. China Phys. Mech. Astron.* **66** (2023) 129811 [[arXiv:2310.03066](#)] [[INSPIRE](#)].
- [45] R. Zhou et al., *DESI luminous red galaxy samples for cross-correlations*, *JCAP* **11** (2023) 097 [[arXiv:2309.06443](#)] [[INSPIRE](#)].
- [46] B.D. Wandelt, E. Hivon and K.M. Górski, *Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology*, *Phys. Rev. D* **64** (2001) 083003 [[astro-ph/0008111](#)] [[INSPIRE](#)].
- [47] D. Alonso, J. Sanchez and A. Slosar, *A unified pseudo- C_ℓ framework*, *Mon. Not. Roy. Astron. Soc.* **484** (2019) 4127 [[arXiv:1809.09603](#)] [[INSPIRE](#)].
- [48] C. García-García, D. Alonso and E. Bellini, *Disconnected pseudo- C_ℓ covariances for projected large-scale structure data*, *JCAP* **11** (2019) 043 [[arXiv:1906.11765](#)] [[INSPIRE](#)].
- [49] N.A. Maksimova et al., *AbacusSummit: a massive set of high-accuracy, high-resolution N-body simulations*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 4017 [[arXiv:2110.11398](#)] [[INSPIRE](#)].
- [50] L.H. Garrison, D.J. Eisenstein and P.A. Pinto, *A High-Fidelity Realization of the Euclid Code Comparison N-body Simulation with Abacus*, *Mon. Not. Roy. Astron. Soc.* **485** (2019) 3370 [[arXiv:1810.02916](#)] [[INSPIRE](#)].
- [51] L.H. Garrison et al., *The abacus cosmological N-body code*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 575 [[arXiv:2110.11392](#)] [[INSPIRE](#)].
- [52] B. Hadzhiyska, L.H. Garrison, D. Eisenstein and S. Bose, *The halo light-cone catalogues of AbacusSummit*, *Mon. Not. Roy. Astron. Soc.* **509** (2022) 2194 [[arXiv:2110.11413](#)] [[INSPIRE](#)].
- [53] R. Zhou et al., *Target Selection and Validation of DESI Luminous Red Galaxies*, *Astron. J.* **165** (2023) 58 [[arXiv:2208.08515](#)] [[INSPIRE](#)].
- [54] S. Yuan et al., *The DESI one-per cent survey: exploring the halo occupation distribution of luminous red galaxies and quasi-stellar objects with AbacusSummit*, *Mon. Not. Roy. Astron. Soc.* **530** (2024) 947 [[arXiv:2306.06314](#)] [[INSPIRE](#)].
- [55] D.N. Limber, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field*, *Astrophys. J.* **117** (1953) 134.
- [56] A.J.S. Hamilton, *Linear redshift distortions: A Review*, in *The Evolving Universe, Astrophysics and Space Science Library* **231**, D. Hamilton ed., Springer, Dordrecht, Netherlands (1998), pp. 185–275 [[DOI:10.1007/978-94-011-4960-0_17](#)] [[astro-ph/9708102](#)] [[INSPIRE](#)].
- [57] K. Tanidis and S. Camera, *Developing a unified pipeline for large-scale structure data analysis with angular power spectra — I. The importance of redshift-space distortions for galaxy number counts*, *Mon. Not. Roy. Astron. Soc.* **489** (2019) 3385 [[arXiv:1902.07226](#)] [[INSPIRE](#)].

- [58] P. McDonald, *Clustering of dark matter tracers: Renormalizing the bias parameters*, *Phys. Rev. D* **74** (2006) 103512 [Erratum *ibid.* **74** (2006) 129901] [[astro-ph/0609413](#)] [[INSPIRE](#)].
- [59] M.M. Ivanov, *Effective Field Theory for Large-Scale Structure*, in *Handbook of Quantum Gravity*, C. Bambi, L. Modesto and I. Shapiro eds., Springer, Singapore (2023) [[DOI:10.1007/978-981-19-3079-9_5-1](#)] [[arXiv:2212.08488](#)] [[INSPIRE](#)].
- [60] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS*, *JCAP* **08** (2009) 020 [[arXiv:0902.0991](#)] [[INSPIRE](#)].
- [61] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rep.* **733** (2018) 1 [[arXiv:1611.09787](#)] [[INSPIRE](#)].
- [62] M.M. Ivanov et al., *Cosmology with the galaxy bispectrum multipoles: Optimal estimation and application to BOSS data*, *Phys. Rev. D* **107** (2023) 083515 [[arXiv:2302.04414](#)] [[INSPIRE](#)].
- [63] J.E. McEwen, X. Fang, C.M. Hirata and J.A. Blazek, *FAST-PT: a novel algorithm to calculate convolution integrals in cosmological perturbation theory*, *JCAP* **09** (2016) 015 [[arXiv:1603.04826](#)] [[INSPIRE](#)].
- [64] X. Fang, J.A. Blazek, J.E. McEwen and C.M. Hirata, *FAST-PT II: an algorithm to calculate convolution integrals of general tensor quantities in cosmological perturbation theory*, *JCAP* **02** (2017) 030 [[arXiv:1609.05978](#)] [[INSPIRE](#)].
- [65] F. Chollet et al., *Keras*, (2015), <https://keras.io>.
- [66] M. Abadi et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, (2015), <https://www.tensorflow.org/>.
- [67] D.P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, in the proceedings of the *International Conference on Learning Representations*, San Diego, CA, U.S.A., 7–9 May 2015, [arXiv:1412.6980](#) [[INSPIRE](#)].
- [68] G. Aricò, R.E. Angulo and M. Zennaro, *Accelerating Large-Scale-Structure data analyses by emulating Boltzmann solvers and Lagrangian Perturbation Theory*, [arXiv:2104.14568](#) [[DOI:10.12688/openreseurope.14310.2](#)] [[INSPIRE](#)].
- [69] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [Erratum *ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [70] J.U. Lange, *nautilus: boosting Bayesian importance nested sampling with deep learning*, *Mon. Not. Roy. Astron. Soc.* **525** (2023) 3181 [[arXiv:2306.16923](#)] [[INSPIRE](#)].
- [71] M.M. Ivanov et al., *Precision analysis of the redshift-space galaxy bispectrum*, *Phys. Rev. D* **105** (2022) 063512 [[arXiv:2110.10161](#)] [[INSPIRE](#)].
- [72] L. Fonseca de la Bella, D. Regan, D. Seery and D. Parkinson, *Impact of bias and redshift-space modelling for the halo power spectrum: Testing the effective field theory of large-scale structure*, *JCAP* **07** (2020) 011 [[arXiv:1805.12394](#)] [[INSPIRE](#)].
- [73] LSST DARK ENERGY SCIENCE collaboration, *Perturbation theory models for LSST-era galaxy clustering: Tests with subpercent mock catalog measurements in Fourier and configuration space*, *Phys. Rev. D* **105** (2022) 123518 [[arXiv:2111.00501](#)] [[INSPIRE](#)].
- [74] LSST DARK ENERGY SCIENCE collaboration, *Galaxy bias in the era of LSST: perturbative bias expansions*, *JCAP* **02** (2024) 015 [[arXiv:2307.03226](#)] [[INSPIRE](#)].
- [75] EUCLID collaboration, *Euclid preparation — XLI. Galaxy power spectrum modelling in real space*, *Astron. Astrophys.* **687** (2024) A216 [[arXiv:2312.00679](#)] [[INSPIRE](#)].

- [76] A. Oddo et al., *Toward a robust inference method for the galaxy bispectrum: likelihood function and model selection*, *JCAP* **03** (2020) 056 [[arXiv:1908.01774](#)] [[INSPIRE](#)].
- [77] G. D’Amico et al., *The BOSS bispectrum analysis at one loop from the Effective Field Theory of Large-Scale Structure*, *JCAP* **05** (2024) 059 [[arXiv:2206.08327](#)] [[INSPIRE](#)].
- [78] A. Chudaykin, M.M. Ivanov and O.H.E. Philcox, *Reanalyzing DESI DR1. I. Λ CDM constraints from the power spectrum and bispectrum*, *Phys. Rev. D* **113** (2026) 063502 [[arXiv:2507.13433](#)] [[INSPIRE](#)].
- [79] T. Lazeyras, C. Wagner, T. Baldauf and F. Schmidt, *Precision measurement of the local bias of dark matter halos*, *JCAP* **02** (2016) 018 [[arXiv:1511.01096](#)] [[INSPIRE](#)].
- [80] T. Lazeyras and F. Schmidt, *Beyond LIMD bias: a measurement of the complete set of third-order halo bias parameters*, *JCAP* **09** (2018) 008 [[arXiv:1712.07531](#)] [[INSPIRE](#)].
- [81] G.S. Farren et al., *Atacama Cosmology Telescope: Multiprobe cosmology with unWISE galaxies and ACT DR6 CMB lensing*, *Phys. Rev. D* **111** (2025) 083516 [[arXiv:2409.02109](#)] [[INSPIRE](#)].
- [82] D. Alonso et al., *Constraining cosmology with the Gaia-unWISE Quasar Catalog and CMB lensing: structure growth*, *JCAP* **11** (2023) 043 [[arXiv:2306.17748](#)] [[INSPIRE](#)].
- [83] DESI collaboration, *DESI DR2 results. II. Measurements of baryon acoustic oscillations and cosmological constraints*, *Phys. Rev. D* **112** (2025) 083515 [[arXiv:2503.14738](#)] [[INSPIRE](#)].
- [84] F.J. Qu et al., *Atacama Cosmology Telescope DR6 and DESI: Structure growth measurements from the cross-correlation of DESI legacy imaging galaxies and CMB lensing from ACT DR6 and Planck PR4*, *Phys. Rev. D* **111** (2025) 103503 [[arXiv:2410.10808](#)] [[INSPIRE](#)].
- [85] A. Barreira, T. Lazeyras and F. Schmidt, *Galaxy bias from forward models: linear and second-order bias of IllustrisTNG galaxies*, *JCAP* **08** (2021) 029 [[arXiv:2105.02876](#)] [[INSPIRE](#)].
- [86] M. Zennaro et al., *Priors on Lagrangian bias parameters from galaxy formation modelling*, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 5443 [[arXiv:2110.05408](#)] [[INSPIRE](#)].
- [87] M.M. Ivanov et al., *The Millennium and Astrid galaxies in effective field theory: comparison with galaxy-halo connection models at the field level*, [arXiv:2412.01888](#) [[INSPIRE](#)].
- [88] M. Gatti et al., *Dark energy survey year 3 results: weak lensing shape catalogue*, *Mon. Not. Roy. Astron. Soc.* **504** (2021) 4312 [[arXiv:2011.03408](#)] [[INSPIRE](#)].
- [89] B. Stölzner et al., *KiDS-Legacy: Consistency of cosmic shear measurements and joint cosmological constraints with external probes*, *Astron. Astrophys.* **702** (2025) A169 [[arXiv:2503.19442](#)] [[INSPIRE](#)].
- [90] T. Matsubara, *Nonlinear perturbation theory with halo bias and redshift-space distortions via the Lagrangian picture*, *Phys. Rev. D* **78** (2008) 083519 [*Erratum ibid.* **78** (2008) 109901] [[arXiv:0807.1733](#)] [[INSPIRE](#)].
- [91] C. Modi, S.-F. Chen and M. White, *Simulations and symmetries*, *Mon. Not. Roy. Astron. Soc.* **492** (2020) 5754 [[arXiv:1910.07097](#)] [[INSPIRE](#)].
- [92] M. Zennaro et al., *The BACCO simulation project: biased tracers in real space*, *Mon. Not. Roy. Astron. Soc.* **524** (2023) 2407 [[arXiv:2101.12187](#)] [[INSPIRE](#)].
- [93] LSST DARK ENERGY SCIENCE collaboration, *Core Cosmology Library: Precision Cosmological Predictions for LSST*, *Astrophys. J. Suppl.* **242** (2019) 2 [[arXiv:1812.05995](#)] [[INSPIRE](#)].