An analysis of parameter compression and Full-Modeling techniques with Velocileptors for DESI 2024 and beyond

To cite this article: M. Maus *et al* JCAP01(2025)138
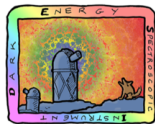
View the article online for updates and enhancements.

DARK ENERGY SPECTROSCOPIC INSTRUMENT (DESI) SURVEY YEAR 1 RESULTS

An analysis of parameter compression and Full-Modeling techniques with Velocileptors for DESI 2024 and beyond

DARK ENERGY
SPECTROSCOPIC
INSTRUMENT

U.S. Department of Energy Office of Science

M. Maus et al.

Full author list at the end of the paper

E-mail: mark.maus@berkeley.edu

ABSTRACT: In anticipation of forthcoming data releases of current and future spectroscopic surveys, we present the validation tests and analysis of systematic effects within `velocileptors` modeling pipeline when fitting mock data from the `AbacusSummit` N-body simulations. We compare the constraints obtained from parameter compression methods to the direct fitting (Full-Modeling) approaches of modeling the galaxy power spectra, and show that the ShapeFit extension to the traditional template method is consistent with the Full-Modeling method within the standard ΛCDM parameter space. We show the dependence on scale cuts when fitting the different redshift bins using the ShapeFit and Full-Modeling methods. We test the ability to jointly fit data from multiple redshift bins as well as joint analysis of the pre-reconstruction power spectrum with the post-reconstruction BAO correlation function signal. We further demonstrate the behavior of the model when opening up the parameter space beyond ΛCDM and also when combining likelihoods with external datasets, namely the Planck CMB priors. Finally, we describe different parametrization options for the galaxy bias, counterterm, and stochastic parameters, and employ the halo model in order to physically motivate suitable priors that are necessary to ensure the stability of the perturbation theory.

Contents

1   Introduction

The large-scale structure (LSS) of the Universe is the observed, coherent spatial distribution of material on scales larger than the typical galaxy or halo scale, and provides a powerful observational tool for probing cosmic evolution. LSS observations allow us to study 3D volumes of the sky that span a long range of cosmic times, enabling us to study the initial conditions of the primordial universe as well as its evolution at later times. [1–4].

One of the primary methods of measuring the evolution of LSS is through galaxy redshift surveys that aim to probe the clustering of matter on a wide range of scales using galaxies as tracers. Spectroscopic galaxy surveys have had significant success over the years in scanning large regions of the sky. These include the 2dF [5], 6dF [6], GAMA [7], WiggleZ [8], and most recently the completed Sloan Digital Sky Survey (SDSS), composed of data from SDSS, SDSS-II [9], BOSS [10–12], and eBOSS [13–15]. The next telescope surveys to further push the boundaries of LSS observations that have recently begun operations are the Euclid Satellite [16, 17] and the ground-based Dark Energy Spectroscopic Instrument (DESI) [18–20]. DESI aims to cover over 14,000 deg^2 by the end of 5 years of observations, with target samples of stars from the Milky Way Survey (MWS), bright galaxies from the Bright Galaxy Survey (BGS, $0.0 < z < 0.4$), Luminous Red Galaxies (LRG, $0.4 < z < 1.1$), Emission Line Galaxies (ELG, $1.1 < z < 1.6$), and Quasars (QSO, $1.6 < z < 2.1$). Altogether the DESI survey will span an effective volume of about $20\,(h^{-1}\mathrm{Gpc})^3$ by the end of its 5 years of observation [21].

In anticipation of the upcoming Year-1 data release of DESI [22–30] (as well as later releases along with Euclid), it is important to characterize the performance of the current state-of-the-art models for analyzing the observed galaxy clustering 2-point statistics and the resultant cosmological constraints. The growth of large-scale structure is a competition between gravity, the dominant force on large scales, and the expansion of the universe. Models must also include several other effects: first, galaxies are not perfect tracers of the underlying matter overdensity field, and thus a 'biasing' scheme is needed in order to relate the matter power spectrum to the observed galaxy spectrum (see ref. [31] for a recent review). Second, since distances along the line-of-sight (LOS) are inferred from redshifts, components of galaxy peculiar velocities in the LOS direction influence the inferred distances and are a source of anisotropy in the observed clustering signal [32, 33]. This latter effect is known as redshift space distortions (RSD) and provides both a challenge to modeling while also giving direct access to information about the growth rate of LSS. Finally, nonlinear effects on small scales must be included. We use perturbation theory to model the mildly non-linear regime, with additional parameters to account for the small-scale physics such that the models are not sensitive to the complicated processes e.g. involved with galaxy formation (sometimes known as Effective Field Theory or EFT terms [34–36]). The model considered in this work, `velocileptors`[1] [37, 38], is one of the models that will be used for analyzing the full-shape power spectra from the upcoming DESI survey data releases, the others being the Fourier space Eulerian PT codes `PyBird` [39–41] and `FOLPS`ν [42] and the configuration space code `EFT-GSM` [43]. The purpose of this work is to characterize the performance of `velocileptors` and understand any systematic issues by comparing to a suite of simulated,

---

[1] https://github.com/sfschen/velocileptors/tree/2.0.

or 'mock', data. Similar tests are being performed with the other three models in addition to a comparison between models, and will be reported in companion publications [44–47]. While `velocileptors` has been tested previously on simulations [38, 48, 49], here we focus on DESI-like galaxies and redshift ranges, and also use the new AbacusSummit [50] suite of simulations produced for the DESI collaboration that is also used to test the other theory models.

Within the framework of the model, there are still various approaches to fitting data. One method, previously used by the BOSS and eBOSS collaborations, involved choosing a fiducial template for the linear power spectrum while compressing the observed power spectrum multipoles into three parameters: the amplitude of the redshift-space anisotropy $f\sigma_8$, and the two scaling parameters parallel and perpendicular to the line of sight, i.e. $\alpha_\parallel$ and $\alpha_\perp$. This technique was meant to encode the intuition that, for currently popular cosmological models, primary CMB anisotropies fix the parameters determining the shape of the power spectrum but late-time effects such as non-trivial dark energy evolution or spatial curvature can affect the total growth and the distance-redshift relation. These impacts are accounted for by the three parameters above and redshift surveys can constrain them well. An extension to this standard "template" fit is to include another compressed "ShapeFit" parameter to allow a set of modifications to the shape of the linear power spectrum [51]. The extra shape information of this method allows for tighter constraints on cosmological parameters when interpreting the compressed statistics in light of a given cosmological model without including CMB priors. This partially bridges the gap in constraining power between the traditional template fit and the direct fitting or "Full-Modeling" approach of directly varying the parameters of a specific cosmological model. In this paper we compare these three methods under a variety of conditions in order to better understand the advantages and disadvantages of the methods. A comparison of the template and Full-Modeling approaches was investigated in ref. [52] on the BOSS DR12 dataset, specifically focusing on shifts in $f\sigma_8$ constraints between the two methods. Here we extend that analysis to include the ShapeFit method and compare the three methods for the range of different settings, parameterizations, and modeling choices.

This paper is organized as follows. We begin by describing the Abacus simulations in section 2 and give an overview of Lagrangian Perturbation Theory (LPT) and `velocileptors` in section 3. We describe the parameter compression and Full-Modeling fitting methods in more detail in section 4. The results of our primary tests, namely the dependence on scale cuts, joint fitting of multiple redshift bins, post-reconstruction statistics, wCDM models, CMB priors, varying n_s, Lagrangian vs Eulerian (EPT) Perturbation Theory, and freeing σ_8 are presented in section 5. We conclude the paper in section 6. We also provide a brief discussion of our method for analytic marginalization over the linear parameters in our model in appendix A along with some further tests, namely the dependence of ω_b prior, inclusion of cubic bias, and inclusion of hexadecapole moment in appendix D. In appendix B we discuss the issue of parameter projection effects and the dependence on priors within our model, a problem that also arises in many other areas of cosmology. We follow this up with a section dedicated to the halo model in appendix C, which allows us to estimate typical scales for stochastic parameters in our model and provide physical motivation for our prior choices. Appendix E explains our use of emulators based on Taylor series in order to speed up likelihood evaluations, and we show that they perform consistently with the direct theory predictions.
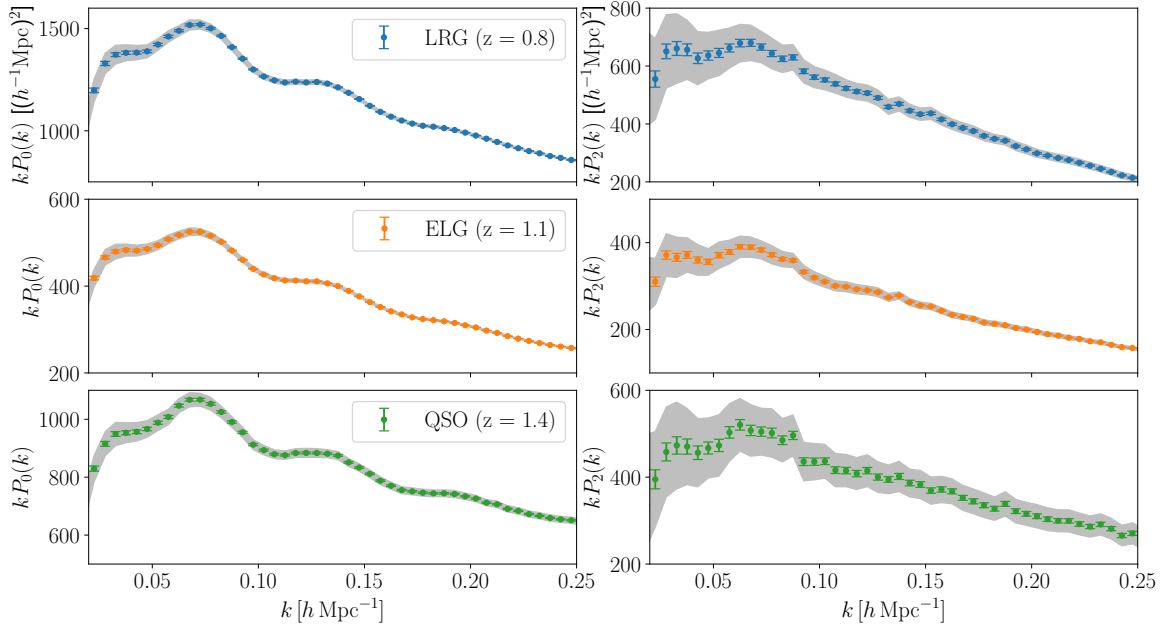
Figure 1. Power spectrum monopole (left) and quadrupole (right) mock data for the LRG, ELG, and QSO tracers. For each tracer, the mean of the 25 N-body realizations is used. The error bars of the data correspond to the covariance re-scaled by the number of realizations, which represents a survey volume of 200 $h^{-3}\mathrm{Gpc}^3$. The shaded regions show the error bars for a single cubic box, of volume $V = 8\,h^{-3}\mathrm{Gpc}^3$.

2   Mock data

To test our theory model we make use of the AbacusSummit [50] suite of N-body simulations in their native, cubic geometry. These simulations were run with the `Abacus` [53] N-body code on the Summit supercomputer at the Oak Ridge Leadership Computing Facility for use by the DESI collaboration. The simulations relevant to this work use a fixed cosmology,[2] with 25 boxes each with a different random number seed for the initial conditions run in a $(2\,h^{-1}\mathrm{Gpc})^3$ volume for a combined volume of $200h^{-3}\mathrm{Gpc}^3$. The mock galaxy catalogs have been produced for three types of tracers, each produced at a different redshift: Luminous Red Galaxies (LRGs) at $z = 0.8$, Emission Line Galaxies (ELGs) at $z = 1.1$, and Quasars (QSOs) at $z = 1.4$.[3] For this study we ignore light-cone and evolution effects in order to better study the non-linear dynamics and biasing models. The RSD power spectrum data for each tracer is shown in figure 1.

---

[2]The Abacus fiducial cosmology has $h = 0.6736$, $\omega_b = 0.02237$, $\omega_{\mathrm{cdm}} = 0.12$, $A_s = 2.0830 \times 10^{-9}$, and $n_s = 0.9649$, with a corresponding BAO drag scale of $r_d = 99.08\,h^{-1}\mathrm{Mpc}$.

[3]The constraining power from a single redshift bin is similar to that expected for each tracer by year-5 of the DESI survey. While the real LRG data will actually be split into multiple redshift bins, the constraints from the joint analyses will be similar to those obtained from the single LRG bin in this work. We do not expect the conclusions in this paper to change significantly if the mocks had been produced in more redshift bins for each tracer. However, projection effects are expected to be more significant in extended models in Year-1 as the data is not as constraining yet as these mocks. This is discussed further in appendix B.1.

The covariance we use for each tracer is calculated by Monte-Carlo from 1000 "effective Zeldovich approximation" (EZmock [54]) simulations of the same cosmology.[4] We compute this covariance numerically via:

$$Cov[P(k)]_{ij} = \frac{1}{N-1} \sum_{n}^{N} [P_n(k_i) - \langle P(k_i) \rangle][P_n(k_j) - \langle P(k_j) \rangle]^\top \qquad (2.1)$$

In principle, when using as data the mean of 25 cubic boxes the error bars of the data should also be re-scaled to reflect the increase in volume because $\sigma^2 \propto V^{-1}$. A proper treatment of the mean of 25 realizations would therefore involve re-scaling the covariance from the EZmocks by a factor of $1/25$. However, we must be careful in interpreting results when the error bars of the data are so tight, as the "survey volume" of the simulations is orders of magnitude larger than any realistic survey will ever be able to achieve. For example, if we consider a future survey covering $18\,000$ deg^2 with tracers in a single redshift bin spanning $0.75 < z < 1.25$, then the comoving volume of that data would be about $24\,(h^{-1}\mathrm{Gpc})^3$, which is still much less than the $200\,(h^{-1}\mathrm{Gpc})^3$ volume of the simulations. The $8(h^{-1}\mathrm{Gpc})^3$ volume of a single box in our simulations is much closer to what we expect for any tracers/redshift bin by the end of five years of DESI observations.

The motivation for the large simulation volume is to detect systematic errors in the models relevant to the DESI Y5 data. If we define the detection of a systematic error as being larger than twice the statistical error σ_{sim} of the simulations and would like to keep systematic errors below some fraction $1/n$ of the Y5 data errors (σ_{Y5}), then this implies that we desire simulations with $\sigma_{\mathrm{sim}} \leq (2n)^{-1}\sigma_{Y5}$. If $\sigma \propto 1/\sqrt{V}$, then for $n = 3$ and a DESI Y5 volume of $5\,(h^{-1}\mathrm{Gpc})^3$, we would require a simulation volume of $180\,(h^{-1}\mathrm{Gpc})^3$. The Abacus simulations fulfill this requirement. However the above argument fails to account for the systematic errors of the N-body simulations themselves. The fractional errors of the Abacus mock LRG monopole data with 25 box covariance (re-scaled by $1/25$) are roughly 0.15% between $0.15 < k < 0.2\,h\mathrm{Mpc}^{-1}$. Ref. [55] compared different cosmological N-body codes and found that RSD power spectra multipoles differed by $\approx 0.5\%$ in the same k-range, i.e. the simulations themselves do not agree to these levels of precision, even before uncertainties from initial condition generation, halo finding and additional physics are included [56]. In addition to this, the large volume also reflects a level of precision that our models are not designed for, meaning that contributions from, e.g., two-loop terms that we don't include in our theory can result in poor fits. For all of these reasons, we will primarily focus on results using the un-rescaled covariance of the more reasonable single-box volume in the analysis of this paper, while only commenting briefly on the 25 box covariance results when relevant. Finally, when computing the covariance from a finite number of simulations, one should in principle include corrections such as the Hartlap factor [57], which depends on the number of bins in the data vector versus the number of independent mock data sets used. Given the large number of EZmock simulations that we use, this factor is close to 1 and we therefore do not observe any noticeable change in constraints when including the correction. We also

---

[4]Since these computationally efficient simulations make use of the Zel'dovich approximation they may not be as accurate at small scales. As we will show later, our models are able to obtain unbiased constraints up to $k_{\mathrm{max}} = 0.2\,h\mathrm{Mpc}^{-1}$ but analytic covariances may be desirable in the future.

do not observe any significant bias in constraints arising from the finite number of mocks and therefore neglect the Hartlap correction in our analyses.

3 Theory and model

The `velocileptors` code is based on the Lagrangian Perturbation Theory (LPT) approach to large-scale structure. This approach treats dark matter as collisionless particles whose mapping from initial (Lagrangian) positions, $\boldsymbol{q}$, to their final observed coordinates, $\boldsymbol{x}$ is given by $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\Psi}(\boldsymbol{q})$, where $\boldsymbol{\Psi}(\boldsymbol{q})$ is the displacement field. The dynamical equation, based on Newtonian gravity in an expanding spacetime, $\ddot{\Psi} + \mathcal{H}\dot{\Psi} = -\nabla_{\boldsymbol{x}}\Phi$, is perturbatively expanded and solved as $\boldsymbol{\Psi} = \boldsymbol{\Psi}^{(1)} + \boldsymbol{\Psi}^{(2)} + \boldsymbol{\Psi}^{(3)} + \dots$. The observed galaxy overdensity is derived from number conservation, with the inclusion of a bias functional in the initial conditions, $F[\delta_0(\boldsymbol{q})]$, that relates the tracer overdensity field to the linear matter field in the form of a Taylor series [37, 38]. In Fourier space, this results in

$$1 + \delta_g(\boldsymbol{k}) = \int d^3\boldsymbol{q} \; F[\delta_0(\boldsymbol{q})] e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\Psi}(\boldsymbol{q}))}$$

$$F[\delta_0(\boldsymbol{q})] = 1 + b_1\delta_0 + \frac{1}{2}b_2(\delta_0(\boldsymbol{q})^2 - \left\langle \delta_0^2 \right\rangle) + b_s(s_0^2(\boldsymbol{q}) - \left\langle s_0^2 \right\rangle) + b_3\mathcal{O}_3(\boldsymbol{q}), \quad (3.1)$$

where $s_0 = (\partial_i\partial_j/\partial^2 - \delta_{ij}/3)\delta_0$ is the initial shear tensor. The Lagrangian biases b_O describe the response of galaxy formation to large-scale perturbations and are the free parameters of the theory — absent a complete model of galaxy formation at small scales their values must be measured directly from large-scale observables like the power spectrum, though rough estimates for their sizes can be made through toy models like halo occupation distributions. At 1-loop order there is only one non-degenerate cubic bias contribution which we include schematically as $\mathcal{O}_3$. Note that the Lagrangian bias parameters here are not equivalent to the Eulerian ones (for example the standard linear bias is $b = 1 + b_1$) but equivalent under a set of linear transformations (see e.g. ref. [38]). Throughout most of this paper we will set $b_3 = 0$ under the assumption that the cubic nonlinearities in galaxy clustering are consistent with those from dynamical contributions alone [58]. We test this assumption in appendix D.

The modeling of observed galaxy clustering statistics is complicated by the peculiar velocities of the galaxies, whose line-of-sight components introduce anisotropies in the clustering signal, an effect known as Redshift Space Distortions (RSD). In LPT, the transformation into redshift space amounts to a boost along the LOS direction, $\hat{n}$ so that the redshift space displacement field is

$$\boldsymbol{\Psi}_s = \boldsymbol{\Psi} + \dot{\boldsymbol{\Psi}} = \boldsymbol{\Psi} + \frac{\hat{n}(\mathbf{v}\cdot\hat{n})}{\mathcal{H}}, \quad (3.2)$$

where $\mathbf{v}$ is the galaxy peculiar velocity and $\mathcal{H}$ is the conformal Hubble parameter. We can simplify this relation with the Einstein-deSitter Approximation (EdS), such that

$$\boldsymbol{\Psi}_s^{(n)} = \boldsymbol{\Psi}^{(n)} + nf(\hat{n}\cdot\boldsymbol{\Psi}^{(n)}), \quad (3.3)$$

where f is the linear growth rate. This can be expressed as a rotation of the real space field via the matrix $R^{(n)} = \delta_{ij} + nf\hat{n}_i\hat{n}_j$ such that $\boldsymbol{\Psi}^{s,(n)} = R^{(n)}\boldsymbol{\Psi}^{(n)}$. Defining the pairwise

displacement field in redshift space as $\Delta_s = \Psi_s(\boldsymbol{q}_1) - \Psi_s(\boldsymbol{q}_2)$, the redshift-space galaxy power spectrum can be obtained from the cumulant expansion of

$$P_{s,g}(\boldsymbol{k}) = \int d^3\boldsymbol{q} \left\langle e^{i\boldsymbol{k}\cdot(\boldsymbol{q}+\Delta_s)} F(\boldsymbol{q}_1)F(\boldsymbol{q}_2) \right\rangle_{\boldsymbol{q}=\boldsymbol{q}_1-\boldsymbol{q}_2}. \tag{3.4}$$

In order to accurately capture the effects of long-wavelength (IR) linear displacements on the power spectrum, particularly with respect to their smearing of the BAO, it is necessary to include their effects beyond 1-loop order in perturbation theory [59–62]. This class of techniques is known in the literature as "IR resummation": in our scheme the linear piece, i.e. the $A_{ij}^{s,(11)}$ component of $A_{ij}^s = \left\langle \Delta_i^s \Delta_j^s \right\rangle$, is split into long- and short- wavelength components, $A_{ij}^{s,\mathrm{lin}} = A_{ij}^{s,<} + A_{ij}^{s,>}$, with a cutoff scale k_{IR}, and we keep the $A_{ij}^{s,<}$ piece exponentiated while expanding all other contributions to 1-loop order. Due to the matrix transformation between the real and redshift space displacements, $\boldsymbol{\Psi}^{s,(n)} = R^{(n)}\boldsymbol{\Psi}^{(n)}$, both velocities and displacements contribute to the resummed A_{ij}^s. The expression for the power spectrum becomes [38]

$$\begin{aligned}
P_{s,g}^{PT}(\boldsymbol{k}) = \int d^3\boldsymbol{q} \; e^{i\boldsymbol{k}\cdot\boldsymbol{q}} e^{-\frac{1}{2}k_ik_jA_{ij}^{s,<}} &\left\{ 1 - \frac{1}{2}k_ik_jA_{ij}^{s,>} + \frac{1}{8}k_ik_jk_kk_lA_{ij}^{s,>}A_{kl}^{s,>} \right. \\
&- \frac{1}{2}k_ik_jA_{ij}^{s,\,\mathrm{loop}} + \frac{i}{6}k_ik_jk_kW_{ijk}^s + 2ib_1k_i\left(1 - \frac{1}{2}k_ik_jA_{ij}^{s,>}\right)U_i^s - b_1k_ik_jA_{ij}^{s,10} \\
&+ b_1^2\left(1 - \frac{1}{2}k_ik_jA_{ij}^{s,>}\right)\xi_{\mathrm{lin}} + ib_1^2k_iU_i^{s,11} - b_1^2k_ik_jU_i^{s,\mathrm{lin}}U_j^{s,\mathrm{lin}} \\
&+ \frac{1}{2}b_2^2\xi_{\mathrm{lin}}^2 + 2ib_1b_2\xi_{\mathrm{lin}}k_iU_i^{s,\mathrm{lin}} - b_2k_ik_jU_i^{s,\mathrm{lin}}U_j^{s,\mathrm{lin}} + ib_2k_iU_i^{s,20} \\
&\left. + b_s\left(-k_ik_j\Upsilon_{ij}^s + 2ik_iV_i^{s,10}\right) + 2ik_ib_1b_sV_i^{s,12} + b_2b_s\chi + b_s^2\zeta + 2ib_3k_iU_{b_3,i}^s + 2b_1b_3\theta + \dots \right\}.
\end{aligned} \tag{3.5}$$

The other correlators appearing above (ξ, W, V, U, etc.) are defined in [37, 38, 59, 63].

We account for the sensitivity to small scales by introducing counterterms with coefficients, α_n, that multiply the tree-level power spectrum. These coefficients describe couplings with short-wavelength modes whose sizes are not directly specified by perturbation theory. While their exact values (or even signs) are not known, we can put reasonable priors on them based on the size of gravitational nonlinearities seen in N-body simulations and expected nonlocalities induced by galaxy formation and baryonic physics, all of which contribute additively to the α_n. Equivalently, the expected contribution of these effects dictates the scales on which our perturbative model is valid. We therefore put Gaussian priors on each counterterm centered at zero with widths set such that their corrections are perturbative at our chosen k_{max}. We similarly include stochastic contributions which we parametrize with $\mathrm{SN}_0 = R_h^3$, $\mathrm{SN}_2 = R_h^3\sigma_2$, and $\mathrm{SN}_4 = R_h^3\sigma_4$, where R_h^3 is the typical galaxy or halo formation scale and the σ_n arise from correlations of stochastic modes in densities and velocities, (e.g. $\langle\delta v\rangle$, $\langle v^2\rangle$, etc.). These stochastic terms again account for the small-scale modes missing in perturbation theory, whose signs and exact values are unknown, but whose rough size can be estimated based on our understanding of the small-scale distribution and velocities of galaxies in halos (see section 4.2 and appendix C and also ref. [64]). These contributions are

added to the 1-loop power spectrum, $P_{s,g}^{PT}(\boldsymbol{k})$, above to give our final LPT prediction

$$P_{s,g}(\boldsymbol{k}) = P_{s,g}^{PT}(\boldsymbol{k}) + (b + f\mu^2)(b\alpha_0 + f\alpha_2\mu^2 + f\alpha_4\mu^4)k^2 P_{s,b_1^2}(\boldsymbol{k})$$
$$+ (\mathrm{SN}_0 + \mathrm{SN}_2 k^2\mu^2 + \mathrm{SN}_4 k^4\mu^4), \tag{3.6}$$

where P_{s,b_1^2} is the term containing $b_1^2\xi_{\mathrm{lin}}$ in eq. (3.5) evaluated to linear order outside of the exponential. This parameterization of the counterterms differs slightly from previous works using `velocileptors`. While giving consistent results, it makes it easier to interpret the counterterms as "fractional corrections" to the linear theory multipoles and motivates our choice of prior width on these parameters. For example, a value of $\alpha_n = 12.5\,h^{-2}\mathrm{Mpc}^2$ corresponds to a 50% correction to the nth moment at $k_{\max} = 0.20\,h\mathrm{Mpc}^{-1}$. We also note that even though this parameterization may appear to introduce new degeneracies within the counterterms, we find no significant change in constraints or increased projection effects.

In computing the observed power spectrum, we assume a fiducial cosmology to convert θ and z to 3D distances using the fiducial distance-redshift relation. We need to account for distortions in $P(k)$ between assumed and true coordinates, the "Alcock-Paczynski (AP) effect" [65], in our modeling. We do this by rescaling the theoretical power spectrum in true cosmological coordinates to the observed coordinates by:

$$P_s^{\mathrm{obs}}(\boldsymbol{k}_{\mathrm{obs}}) = q_\perp^{-2} q_\parallel^{-1} P_s(\boldsymbol{k}), \quad k_{\parallel,\perp}^{\mathrm{obs}} = q_{\parallel,\perp}\,k_{\parallel,\perp}, \tag{3.7}$$

with the scaling parameters above are defined by:[5]

$$q_\parallel = \frac{H^{\mathrm{ref}}(z)}{H(z)}, \quad q_\perp = \frac{D_A(z)}{D_A^{\mathrm{ref}}(z)}. \tag{3.8}$$

$D_A(z)$ is the comoving angular diameter distance and the "ref" superscript labels the values from the fiducial cosmology.

Finally, we use a Legendre transformation to compute the predicted power spectrum as multipoles,

$$P_\ell(k_{\mathrm{obs}}) = \frac{(2\ell+1)}{2} \int_{-1}^1 d\mu\ P(k,\mu_{\mathrm{obs}})\mathcal{L}_\ell(\mu) \tag{3.9}$$

where $\mathcal{L}_\ell(\mu)$ is the Legendre polynomial of order ℓ.

4 Fitting methods

4.1 Standard template and ShapeFit

The traditional parameter compression method used originally by the BOSS/eBOSS collaborations involves choosing a reference cosmology, $\boldsymbol{\Theta}^{\mathrm{ref}}$, and keeping the resultant linear power spectrum, and by extension, the dependence on early-universe physics, fixed. The "compressed" parameters being varied are then the amplitude, $f\sigma_{s8}$ and the distance scalings transverse and along the line-of-sight, $\alpha_\perp, \alpha_\parallel$; all of which are only dependent on late-time

---

[5]Previously in BOSS analyses(e.g. [48, 52]) we have used the notation $\alpha_{\parallel,\perp}, \tilde{\alpha}_{\parallel,\perp}$ in place of $q, \alpha_{\parallel,\perp}$ but in this paper we use the latter in order to be consistent with the conventions of other DESI papers.

dynamics. The quantity $f\sigma_{s8}$, which controls the ratio of monopole-to-quadrupole amplitudes, is a product of the growth rate, $f \simeq \Omega_m^{0.55}$ and the total amplitude, σ_{s8}, at $R = s \cdot 8\,h^{-1}\mathrm{Mpc}$ scales. Here $s = r_\mathrm{d}/r_\mathrm{d}^\mathrm{fid}$ with r_d being the BAO scale at the drag epoch. We will comment on the s scaling further below. The two distance scaling parameters are defined by,

$$\alpha_\parallel = \frac{H^\mathrm{ref}(z)}{H(z)}\left(\frac{r_\mathrm{d}^\mathrm{ref}}{r_\mathrm{d}}\right) = q_\parallel \left(\frac{r_\mathrm{d}^\mathrm{ref}}{r_\mathrm{d}}\right) = \frac{q_\parallel}{s}, \quad \alpha_\perp = \frac{D_A(z)}{D_A^\mathrm{ref}(z)}\left(\frac{r_\mathrm{d}^\mathrm{ref}}{r_\mathrm{d}}\right) = q_\perp \left(\frac{r_\mathrm{d}^\mathrm{ref}}{r_\mathrm{d}}\right) = \frac{q_\perp}{s}. \quad (4.1)$$

We highlight that these parameters used in the template fitting are different from the scaling parameters defined in eq. (3.8) by a factor of $\left(r_\mathrm{d}^\mathrm{ref}/r_\mathrm{d}\right)$.[6] This is because in the template method we assume that most information comes from the BAO feature, and thus we account for the fact that both changes in r_d and $q_{\parallel,\perp}$ induce stretching in the observed BAO signal.[7] In contrast, with a fitting method in which the underlying cosmology is directly being varied (see next subsection), the changes to r_d affecting the BAO signal are automatically included in the linear power spectrum which is self-consistently varied. We must also emphasize that by including the factors of s in our α scaling parameters we are implicitly assuming distances in units of the BAO scale, which motivates our use of the notation $f\sigma_{s8}$. This subtlety is discussed in detail in section 3 of ref. [51].

Despite sacrificing constraining power through the lack of sensitivity to the early universe (the shape of the transfer function is held fixed by the reference cosmology), this "template" fitting method was sufficient at a time when the tightest constraints on early-time physics came from the CMB and LSS data was too noisy for direct fitting methods to be feasible without significant priors from Planck. The advantages of the template fitting method include the model-independence that allows for mapping the compressed parameter constraints to a cosmological model of one's choosing. Furthermore, computing the linear power spectrum using a Boltzmann code such as `CLASS` or `CAMB` at every step of a Markov Chain Monte Carlo (MCMC) sampler, in addition to calculating nonlinear perturbation theory (PT) corrections, is computationally very expensive. Fixing the linear power spectrum avoids this step, allowing for a faster fitting procedure without needing to train an emulator.

The "ShapeFit" method is an extension to the standard template-fit compression, and was conceived as a way to partially bridge the gap in constraining power between the standard template and direct/full modeling methods, while preserving some of the model-independence of the former technique [51]. This is achieved by allowing modifications to the shape of the linear power spectrum via a multiplicative factor,

$$P'_\mathrm{lin}(\boldsymbol{k}) = P_\mathrm{lin}^\mathrm{ref}(\boldsymbol{k})\ \exp\left\{\frac{m}{a}\tanh\left[a\ln\left(\frac{k}{k_p}\right)\right] + n\ln\left(\frac{k}{k_p}\right)\right\}, \quad (4.2)$$

where $P_\mathrm{lin}^\mathrm{ref}(\boldsymbol{k})$ is the template power spectrum produced by `CLASS` and is fixed throughout the fit. The form of this scaling was an ansatz chosen to best replicate the effect of varying

---

[6]Technically, this "ref" is not necessarily the same as the "ref" in the definitions of $q_{\parallel,\perp}$. The one in $\left(r_\mathrm{d}^\mathrm{ref}/r_\mathrm{d}\right)$ refers to the reference template used in the standard template and ShapeFit fits, whereas in $q_{\parallel,\perp}$ it refers to the fiducial cosmology assumed when converting angles and redshift coordinates to physical distances when measuring the power spectrum. However, in practice it is simplest to choose the same cosmology for the template as was used for measuring the power spectrum from the data, so this distinction is not important.

[7]See discussion in appendix C of ref. [48], where however the pure AP parameters are referred to as α and the BAO-rescaled ones are called $\tilde{\alpha}$.

ω_b, ω_m, and n_s on the shape of the power spectrum (logarithmic slope and small/large scale limits), which would otherwise be captured in the transfer function when running `CLASS`. The modified power spectrum $P'_{\rm lin}(\mathbf{k})$ is what we provide to `velocileptors` to produce the full 1-loop prediction for a given $(f\sigma_8, \alpha_\parallel, \alpha_\perp, m)$. For simplicity we keep fixed the second shape parameter, $n = 0$. Allowing this parameter to vary accounts variations of the template emulating a spectral index effect, which in this paper we do not consider. Following the original ShapeFit paper [51] we choose for a and k_p their proposed values, $a = 0.6$ and $k_p = 0.03\,h\,{\rm Mpc}^{-1}$. With this modification to the classic template analysis, ShapeFit is now able to capture more information from the early universe without sacrificing its model independence. As a drawback, the freedom given by the ShapeFit parametrization in the linear power spectrum may not be sufficient to reproduce the exact shape of the transfer function as modeled by the Direct/Full-Modeling Fit technique (see next subsection) when 1) the fiducial cosmology is very different from the true cosmology, and 2) when the statistical errors of the data are very small. In ref. [44] (figure 2) this effect is quantified for the power spectrum, as well as in an upcoming paper (ref. [66], in prep) focused on DESI Y1 geometry. On another hand, this effect could also be important if the ShapeFit compression technique is applied to higher-order statistics, such as the bispectrum, but this has not been yet quantified, as it goes beyond the scope of this paper.

4.2 Full modeling: ΛCDM and extensions

The alternative modeling technique to parameter compression is a more conventional forward-modeling approach that involves directly varying the underlying parameters of a cosmological model and making a theoretical prediction for the observed quantities. While the ΛCDM model depends on six parameters: $(\omega_b, \omega_{\rm cdm}, H_0, \log(10^{10}A_s), M_\nu$ and $n_s)$, some of these parameters are not constrained by galaxy clustering analyses independently. For these quantities we use priors derived from e.g. Big-Bang Nucleosynthesis (BBN) and/or CMB anisotropies. We initially fix the spectral tilt and neutrino mass to the Abacus fiducial values of $(n_s, M_\nu) = (0.9649, 0.06)$ — though see section 5.7. For the baryon abundance we adopt a narrow gaussian BBN prior of $\mathcal{N}[\mu = 0.02237, \sigma = 0.00037]$ [67] (though see discussion in appendix D). Within these constraints, in this "Full-Modeling" approach the shape of the linear power spectrum is able to change at each step of the MCMC as the shape of the transfer function is dependent on the ΛCDM parameters being varied. If done directly, this method is more computationally expensive because the linear power spectrum must be calculated using a Boltzmann code such as `CLASS` or `CAMB` in addition to the `Velocileptors` PT corrections. However, through the use of an emulator we can efficiently and accurately approximate the predictions for a given set of ΛCDM parameters. Under the assumption that the predicted power spectrum multipoles are a smooth function of the underlying parameters when close to some reasonably chosen values, we can use an emulator based on a Taylor series expansion in the relevant parameter space [41, 48].[8] We find that the emulator agrees well with the direct LPT prediction when going to fourth order in the Taylor expansion. After employing such an emulator both for the Full-Modeling and template/ShapeFit methods,

---

[8]In the event that the data require a significantly different parameter space the analysis can be iterated with the Taylor series recomputed closer to the best fit, assuming the data are sufficiently constraining.

the MCMC chains converge (Gelman-Rubin $|R - 1| < 0.01$) within roughly 1–2 hours.[9] By analytically marginalizing of stochastic and counterterm contributions (see appendix A), the MCMC converges in 5–10 minutes for all methods. Therefore, the improved computational efficiency of a compression is no longer relevant in our setup.

The advantage of the Full-Modeling approach is that it is sensitive to both the early-universe physics that determines the shape of the transfer function, as well as late-time dynamics/geometry. Parameters such as ω_b, ω_{cdm}, and H_0 affect both the early- and late-universe dynamics, and are thus expected to be more tightly constrained in the Full-Modeling approach, when compared to the methods employing a template that fixes the early-universe dependence. On the other hand, the Full-Modeling approach requires choosing a specific cosmological model from the start, and a new MCMC fit is needed for any other model being employed. The parameter compression methods, however, only require one fit, and afterwards the results can be reused and mapped to any model of choice, though the model of choice must be sufficiently close to the template cosmology unlike in the Full-Modeling approach which does not suffer from this requirement.

We show in table 1 the parameters and priors used for the Full-Modeling and ShapeFit methods. We show the priors on bias parameters for three parametrizations. The standard setting in this paper is the "intermediate" freedom case for which the cubic bias is fixed to zero while $(1 + b_1)\sigma_8$, $b_2\sigma_8^2$, and $b_s\sigma_8^2$ are varied with Gaussian priors applied to the latter two. The other parameter choices are discussed in appendix D. We analytically marginalize over the parameters controlling the stochastic and counterterm contributions, and refer readers to appendix A for further details and validation of this method.

Finally we remark that in order to make contact with earlier work, and in particular with our companion papers, we use $\log(10^{10}A_s)$ as the "normalization" of the power spectrum throughout. This choice, being the normalization of the curvature power spectrum at $k = 0.05\,\mathrm{Mpc}^{-1}$, is actually better motivated for CMB surveys than galaxy redshift surveys. Most of the constraining power of our data comes from quasi-linear scales and we better constrain the matter power spectrum than the curvature (or potential) power spectrum. In this respect a better choice for normalization may be σ_8. We will discuss constraints on σ_8 later. We also reiterate that the Full-modeling method does not require any re-scaling of distances by $s = r_d/r_d^{ref}$, and therefore the amplitude being constrained here is σ_8 not σ_{s8}.

4.3 Cosmological inference from compressed statistics

In order to interpret the ShapeFit and standard template results, we must do so in the context of a chosen cosmological model such as ΛCDM. While it is simple to take a set of ΛCDM parameters and compute the distances, $H(z)$, $D_A(z)$, and r_d using CLASS or CAMB, in order to compute compressed parameters assuming a certain fiducial cosmology, it is more tricky in reverse [51]. Instead we must fit ΛCDM parameters to the results of a fixed template fit with another MCMC. We take the chains in the compressed parameters that were obtained from the initial template fits, and compute the parameter mean vector and covariance matrix, i.e. $\bar{\Theta} = (f\bar{\sigma}_8, \bar{\alpha}_{\parallel}, \bar{\alpha}_{\perp}, \bar{m})$ and $\mathbf{C}_{4\times 4}$. Treating $\bar{\Theta}$ and $\mathbf{C}_{4\times 4}$ as a "data" vector and associated covariance, we can now sample in ΛCDM parameters so that for each proposed set of (ω_b,

---

[9]This is when using 8 parallel chains on a single node.

Full-Modeling	ShapeFit	Bias			Stoch/Counter
		Min. F.	Int. F.*	Max. F.	
H_0	$f\sigma_8$	$(1+b_1)\sigma_8$			$\tilde{\alpha}_0$
$\mathcal{U}[55,79]$	$\mathcal{U}[0,2]$	$\mathcal{U}[0.5,3.0]$			$\mathcal{N}[0,12.5]$
$\omega_{\rm b}$	$\alpha_\parallel$	$b_2\sigma_8^2$			$\tilde{\alpha}_2$
$\mathcal{N}[0.02237,0.00037]$	$\mathcal{U}[0.5,1.5]$	$\mathcal{N}[0,5]$	$\mathcal{N}[0,5]$	$\mathcal{N}[0,5]$	$\mathcal{N}[0,12.5]$
$\omega_{\rm cdm}$	$\alpha_\perp$	$b_s\sigma_8^2$			SN_0
$\mathcal{U}[0.08,0.16]$	$\mathcal{U}[0.5,1.5]$	0	$\mathcal{N}[0,5]$	$\mathcal{N}[0,5]$	$\mathcal{N}[0,\mathcal{O}(1/\bar{n}_g)]$
$\log(10^{10}A_{\rm s})$	m	$b_3\sigma_8^3$			SN_2
$\mathcal{U}[2.03,4.03]$	$\mathcal{U}[-3.0,3.0]$	0	0	$\mathcal{N}[0,5]$	$\mathcal{N}[0,\mathcal{O}(f_{\rm sat}\sigma_v^2/\bar{n}_g)]$

Table 1. Velocileptors LPT priors on parameters used in the Full-Modeling (ΛCDM) and ShapeFit fitting methods. The ΛCDM model involves H_0, $\Omega_{\rm b}$, $\omega_{\rm cdm}$, $\log(10^{10}A_{\rm s})$ and all of the bias, stochastic, and counterterms. The ShapeFit method fits $f\sigma_8$, $\alpha_\parallel$, $\alpha_\perp$, m as well as the same bias, stochastic and counterterms. The entries $\mathcal{U}[\min,\max]$ and $\mathcal{N}[\mu,\sigma]$ refer to uniform and Gaussian normal distributions, respectively. For the bias terms we show both minimal, intermediate (standard), and maximal freedom cases, defined in appendix D. For the two counterterms we report the priors within the parameterization for which the counterterms scale relative to the linear theory multipoles. The priors on the stochastic terms are given in table 2 and discussed in the text.

Tracer	$z_{\rm eff}$	$1/\bar{n}_g$	$f_{\rm sat}$	$\log_{10}\bar{M}_h$	$\sigma_v^{\rm est.}$	SN_0	SN_2	SN_4
LRG	0.8	1000	0.1	13.3	7.8	2000	5.0×10^4	1.0×10^6
ELG	1.1	300	0.1	11.9	2.9	1000	2500	2.5×10^4
QSO	1.4	8000	0.03	12.7	5.7	1.5×10^4	5.0×10^4	1.0×10^6

Table 2. Relevant quantities used for the prior widths of stochastic parameters (see text). The typical halo mass, $\log_{10}\bar{M}_h$, per galaxy is expressed in units of $h^{-1}M_\odot$ and $1/\bar{n}_g$ is expressed in $h^{-3}\,{\rm Mpc}^3$. The characteristic velocities, $\sigma_v^{\rm est.}$ are in $h^{-1}{\rm Mpc}$. Motivated by these numbers, the last three cloumns show the widths of the Gaussian priors (centered on 0 and in $h^{-1}{\rm Mpc}$ units) that are used in this paper for each stochastic parameter within each redshift bin. The results do not depend upon the precise values chosen.

$\omega_{\rm cdm}$, h, $\log A_s$) we compute the corresponding vector $\mathbf{\Theta_{thy}} = (f\sigma_8, \alpha_\parallel, \alpha_\perp, m)_{\rm thy}$. Assuming all compressed parameters are Gaussian, we then use an MCMC to sample from the likelihood,

$$\mathcal{L} \propto \exp\left\{-\frac{1}{2}(\mathbf{\Theta_{thy}} - \bar{\mathbf{\Theta}})^T \mathbf{C}_{4\times4}^{-1}(\mathbf{\Theta_{thy}} - \bar{\mathbf{\Theta}})\right\}. \tag{4.3}$$

When inferring cosmological constraints from the ShapeFit parameters, care must be taken in interpreting the amplitude $f\sigma_{s8}$ appropriately, as the slope rescaling via the m parameter also changes σ_{s8}. As noted in refs. [51, 68], the parameter f that is varied in ShapeFit analyses is actually $fA \equiv f(A_{sp}/A_{sp}^{\rm ref})^{1/2}$, where $A_{sp} = s^{-3}P_{\rm no-wiggle}^{\rm lin}(k_p/s,\mathbf{\Theta})$ is the amplitude of the no-wiggle power spectrum at the pivot scale, $k_p \simeq 0.03h{\rm Mpc}^{-1}$. The parameter s describes the scaling of lengths relative to the BAO and is defined to be the ratio $r_{\rm d}/r_{\rm d}^{\rm ref}$. In order to generate the model 1-loop power spectrum multipoles, we must provide

`velocileptors` with the linear power spectrum $P'_{\rm lin}(\boldsymbol{k})$ from eq. (4.2) and the growth factor f. Defining `LPT_RSD` as the function that produces the power spectrum multipoles, the nearly exact degeneracy between f and the power spectrum amplitude (see section 5.9) implies that

$$\texttt{LPT\_RSD}\left[f \times \left(\frac{A_{sp}}{A_{sp}^{\rm ref}}\right)^{1/2}; P'_{\rm lin}(\boldsymbol{k})\right] \leftrightarrow \texttt{LPT\_RSD}\left[f; \left(\frac{A_{sp}}{A_{sp}^{\rm ref}}\right) \times P'_{\rm lin}(\boldsymbol{k})\right], \qquad (4.4)$$

and thus the true $f\sigma_{s8}$ is given by

$$f\sigma_{s8} = f \times \left[\left(\frac{A_{sp}}{A_{sp}^{\rm ref}}\right) \int \frac{dk}{2\pi^2} k^2 \tilde{W}_R^2(kR) P'_{\rm lin}(\boldsymbol{k})\right]^{1/2}$$

$$= f \times \left(\frac{A_{sp}}{A_{sp}^{\rm ref}}\right)^{1/2} \left[\int \frac{dk}{2\pi^2} k^2 \tilde{W}_R^2(kR) P_{\rm lin}(\boldsymbol{k}) \exp\left\{\frac{m}{a} \tanh\left[a \ln\left(\frac{k}{k_p}\right)\right]\right\}\right]^{1/2} \qquad (4.5)$$

$$\simeq \frac{(f\sigma_{s8})^{\rm ref}}{(fA_{sp}^{1/2})^{\rm ref}} f A_{sp}^{1/2} \times \exp\left\{\frac{m}{2a} \tanh\left[a \ln\left(\frac{r_d^{\rm fid}}{R}\right)\right]\right\}. \qquad (4.6)$$

Here R is the smoothing scale of the amplitude parameter σ_R and is chosen to be $R = 8\,h^{-1}{\rm Mpc}$ by convention. There are now two ways in which one could use ShapeFit chains in order to infer about cosmological parameters: one can use the above equations (either the exact or approximate forms) to transform the sampled fA chain into $f\sigma_{s8}$, and then use `CLASS` to compute $f\sigma_{s8}$ for every set of ΛCDM parameters at the interpretation step; or one can directly perform the interpretation on fA by always computing f and A_{sp} while sampling in ΛCDM parameters. We find that the two approaches give consistent constraints in the ΛCDM parameter space.

Finally, the m parameter in ShapeFit that controls the shape of the linear power spectrum can be computed from ΛCDM parameters through the ratio [51]

$$m = \frac{d}{dk}\left(\ln\left[\frac{T(k_p/s, \boldsymbol{\Theta})}{T(k_p, \boldsymbol{\Theta}^{\rm ref})}\right]\right)\Bigg|_{k=k_p}, \quad T(\boldsymbol{\Theta}, k) = \frac{P_{\rm no-wiggle}^{\rm lin}(k, \boldsymbol{\Theta})}{\mathcal{P}_{\mathcal{R}}(k, \boldsymbol{\Theta})}, \qquad (4.7)$$

with primordial power spectrum $\mathcal{P}_{\mathcal{R}}$.

5 Results

Before we present the results from the various systematic tests of `velocileptors` and the different modeling methods, we first revisit the issue of covariance volume. In figure 2 we present 1D posterior constraints from the Full-Modeling fit to LRG mock data as a function of covariance volume, i.e. multiples of the single-box volume such that the covariance is rescaled by $1/n$, $n = 1, 3, 5, \cdots, 15$. We show results for fits using two different k-ranges, $0.02 \le k\,[\,h{\rm Mpc}^{-1}] \le 0.18$ and $0.02 \le k\,[\,h{\rm Mpc}^{-1}] \le 0.20$ (which will be our 'standard' range). We find that as the volume is increased, the constraints in $\Omega_{\rm m}$ shift towards the truth as the error bars tighten, which is indicative of a prior volume effect. For H_0 and $\log(10^{10}A_s)$ the constraints remain mostly stable as the volume is increased, with small shifts increasing with volume that likely relate to the increasing sensitivity to two-loop effects that are not included in the model. For similar reasons, we observe a divergence in constraints between
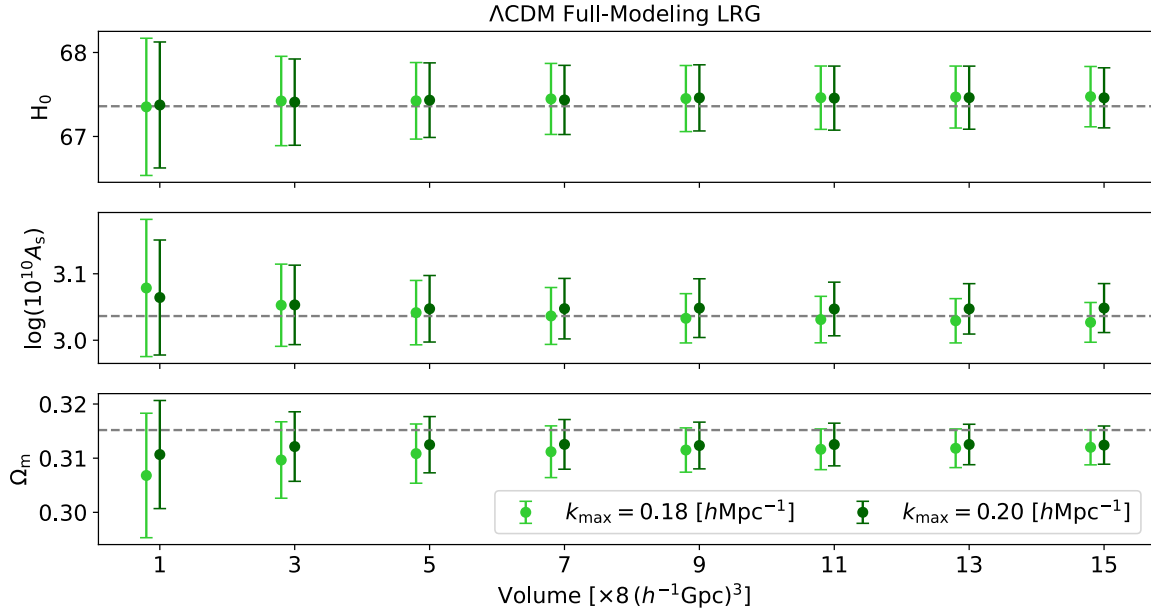
Figure 2. 1D posteriors from the Full-Modeling fit as the covariance volume is varied from that of a single box ($8\,[\,h^{-1}\mathrm{Gpc}]^3$) to 15 boxes ($(120\,[\,h^{-1}\mathrm{Gpc})^3]$).

$k_{\mathrm{max}} = 0.18$ and $k_{\mathrm{max}} = 0.20\,h\mathrm{Mpc}^{-1}$ that grows as the volume is increased. This shows that when using an ultra-tight covariance such as that of the $200\,h^{-1}\mathrm{Gpc}$ simulation volume, one can expect $1 - 2\sigma$ offsets in constraints arising purely from theoretical errors due to the limited number of terms included in the 1-loop power spectrum model. In addition, as mentioned earlier, the N-body simulations themselves have systematic errors that become important at these volumes and can contribute to the shifts we observe.

5.1 Baseline comparison

We begin with comparing constraints in the compressed parameter space between the standard template and ShapeFit approaches, using the single-box covariance, as shown in the left panel of figure 3. We see that the posterior means of the two methods agree very closely, with slightly smaller contours for the standard template due to varying fewer parameters. Since the reference template used in these fits is the true Abacus cosmology, we expect $\alpha_{\parallel,\perp} = 1$ and $m = 0$. In both cases, the means of all parameters are within 1σ of the expected values. When interpreting these results in terms of a ΛCDM cosmology, however, we see a significant difference in the constraints from the two compression methods (right panel of figure 3). While both methods give unbiased constraints on ΛCDM parameters (within 1σ of truth) the error bars for all parameters are significantly larger for the template case due to the lack of information from the power spectrum shape in the template approach. This is expected, as the template method was traditionally combined with external data sets under the assumption that the parameters determining the shape are not as well constrained from LSS data than e.g. CMB anisotropies, but in our setup we rely purely on the LSS data alone(but see section 5.6). Meanwhile, when comparing the constraints between the ShapeFit and Full-Modeling methods, we find a very close agreement in the shape and orientations of
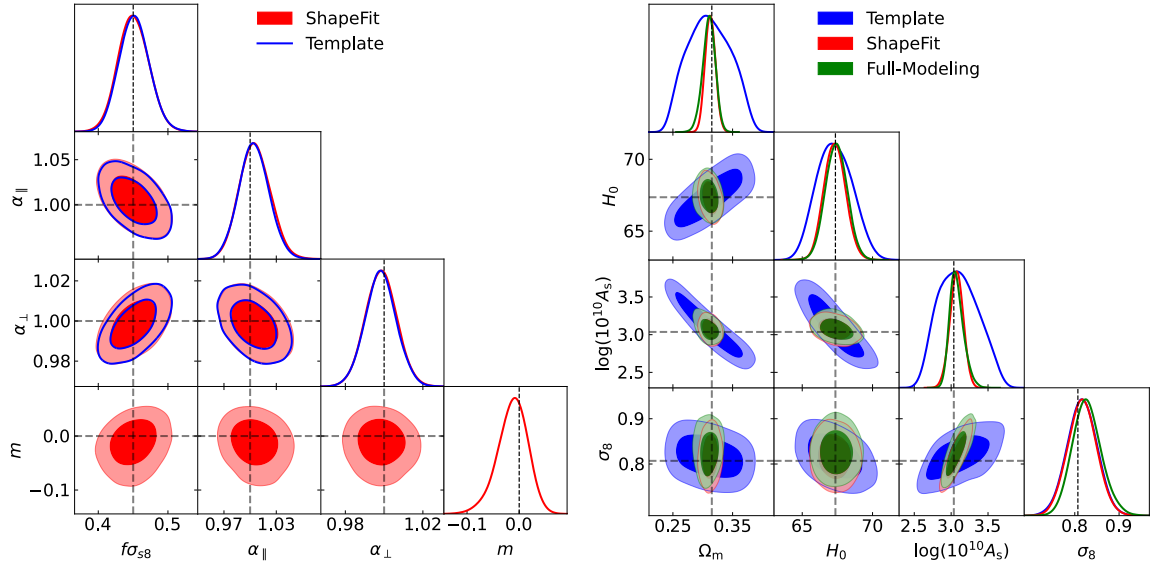
Figure 3. *(left):* comparison of constraints of compressed parameters for the standard template method vs. ShapeFit. *(right):* comparison of constraints on ΛCDM parameters for the standard template, ShapeFit, and full modeling methods. The single-box covariance is used for these results, with our 'standard' $k_{\max} = 0.2\,h\,\mathrm{Mpc}^{-1}$ (see section 5.2 for the discussion of $k_{\max}$ dependence).

the contours, showing that the ShapeFit method is able to match the constraining power of direct model fitting, at least for the ΛCDM case for which it was designed. We do observe mild differences in the tightness of constraints between the ShapeFit and Full-Modeling methods. These could be due to a combination of various approximations in the ShapeFit method, such as controlling the shape of P_{lin} with only one parameter and assuming the compressed parameters to be perfectly Gaussian in the interpretation step.

5.2 Dependence on $k_{\max}$

We next test the dependence on scale cuts of our model, for the different methods. In all cases we fix the lower bound of the k-range to $0.02h\mathrm{Mpc}^{-1}$. This is fully in the linear regime so the stability of the theory is not affected by the specific value chosen, but this choice simply removes points too close to the fundamental mode of the cubic box ($k = 0.003\ h\mathrm{Mpc}^{-1}$). We then run our fits with upper bounds of $k_{\max} = 0.16 - 0.26\,h\mathrm{Mpc}^{-1}$. The results are shown for Full-Modeling and ShapeFit in figure 4 for the LRG, ELG, and QSO tracers. The higher k-modes, above $\sim 0.2\,h\,\mathrm{Mpc}^{-1}$, correspond to smaller scales which are more sensitive to nonlinear effects and galaxy/halo formation physics, which are not well-understood and therefore difficult to model. Our model includes non-linearities only at the 1-loop level and bias only up to cubic order. We therefore expect biases to worsen as higher k-modes are included in the fit. For the single-box volume we find the two methods to remain relatively stable as $k_{\max}$ is increased as the observational errors match or exceed the theoretical or modeling errors, however we do observe $\gtrsim 1\sigma$ offsets in the σ_8 constraints for LRG and ELG tracers in the Full-modeling method when $k_{\max} \geq 0.22\ h\mathrm{Mpc}^{-1}$. We additionally find that for the ELG sample we get more of an tightening of constraints in many parameters
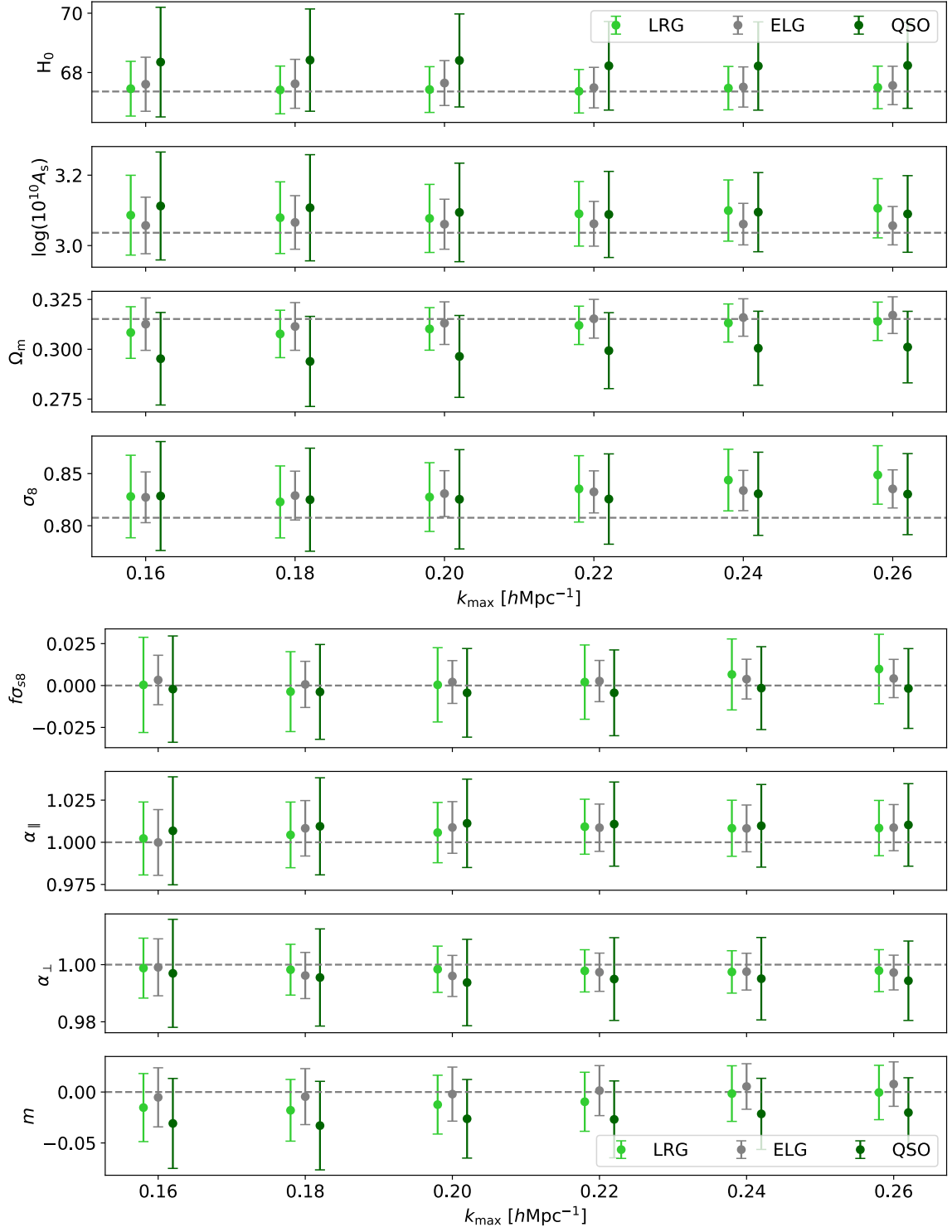
Figure 4. Dependence on $k_{\max}$ of the Full-Modeling (upper panel) and ShapeFit (lower panel) methods for the three tracer types: LRG ($z = 0.8$). ELG ($z = 1.1$), and QSO ($z = 1.4$) (points slightly offset for clarity). The single-box covariance is used for all of these fits.

as $k_{\max}$ is increased than for the other samples. This could be due to the redshift coverage and higher number density of the mock ELG sample.

In figure 5 we repeat this test for the LRG tracers but using the 25 box covariance. We show constraints in the ΛCDM as well as ShapeFit parameter spaces. In this case we obtain significantly biased constraints when $k_{\max} > 0.2 \, h\mathrm{Mpc}^{-1}$. In the ΛCDM parameters, we find a mild improvement in constraining power of Full-Modeling at $\mathrm{k}_{\max} = 0.18 \, h\mathrm{Mpc}^{-1}$ versus our usual setting of $k_{\max} = 0.20 \, h\mathrm{Mpc}^{-1}$. This worsening of constraints when $k_{\max}$ is increased is likely due to a sensitivity to higher-order effects that our theory does not adequately describe, and which become increasingly important with increasing k. When using an extremely tight covariance, the additional high-k points push the fit towards incorrect models and away from the constraints coming from low-k data points. In the compressed parameter space we observe slightly more significant offsets ($\gtrsim 1.\sigma$) in the $\alpha_\perp$ and $f\sigma_{s8}$ constraints for ShapeFit at $\mathrm{k}_{\max} = 0.18 \, h\mathrm{Mpc}^{-1}$. When deriving summary statistics from the Full-Modeling constraints, the $\alpha_\parallel$, and $\alpha_\perp$ parameters are significantly more tightly constrained than in the ShapeFit and Template methods because the ΛCDM priors in Full-Modeling restrict the allowable values that the scaling parameters can take [52]. We use the results from figures 4–5 to motivate a choice of $k_{\max} = 0.20 \, h\mathrm{Mpc}^{-1}$ as our baseline analysis setting, as this is the largest $k_{\max}$ for which all three modeling methods are acceptably close to truth ($\lesssim 1\sigma$ offsets) in the ΛCDM parameter space in both the single-box and full covariance volume cases.

As we proceed to the remainder of tests presented in this paper, we refer readers to figure 6 for a summary figure of 1D constraints on Ω_{m}, H_0, and $\log 10^{10} A_s$ obtained from each of the tests.

5.3 Joint fitting of LRG, ELG, and QSO mocks

We now turn to the joint fitting of data samples from different tracers and redshift bins. The three tracers are Luminous Red Galaxies (LRG, $z = 0.8$), Emission Line Galaxies (ELG, $z = 1.1$), and Quasars (QSO, $z = 1.4$). For the Full-Modeling case, we still sample in ΛCDM parameters as usual but compute separate $P_\ell(k)$ models for each redshift bin and the likelihood is computed from all data sets, i.e. the data vector becomes $\boldsymbol{d} = (P_0^{\mathrm{LRG}}, P_2^{\mathrm{LRG}}, P_0^{\mathrm{ELG}}, P_2^{\mathrm{ELG}}, P_0^{\mathrm{QSO}}, P_2^{\mathrm{QSO}})$. This results in a total effective volume of 600 $(h^{-1}\mathrm{Gpc})^3$. We do not assume any correlation between tracers at different redshifts,[10] so the total joint covariance matrix has zeros in the indices corresponding to cross correlations between different tracers. This ensures that contributions to the log-likelihood such as $\Delta P_i^{\mathrm{LRG}} C_{ij}^{-1} \Delta P_j^{\mathrm{ELG}} = 0$. We use a separate set of nuisance parameters for each type of tracer. For the standard template and ShapeFit fits, the free parameters $(f\sigma_{s8}, \alpha_{\parallel,\perp}, m)$ are in general redshift dependent. While in principle one could use a single $f\sigma_{s8}(z = 0)$ as a free parameter and then rescale by the fiducial growth factor $D(z, \Omega_m = \Omega_m^{\mathrm{fid}})$ in order to get the corresponding parameter for the different samples, the redshift dependence of the α's and m parameters is not as obvious. Instead we perform the parameter compression separately for the LRG, ELG, and QSO samples and obtain three sets of $(f\sigma_{s8}, \alpha_{\parallel,\perp}, m)_z$ to be used

---

[10]The mean data vectors for the LRG, ELG, and QSO tracers actually came from the same 25 realizations and therefore share initial conditions. In principle this means that the redshift bins are not truly uncorrelated, but we assume so in this work for simplicity.
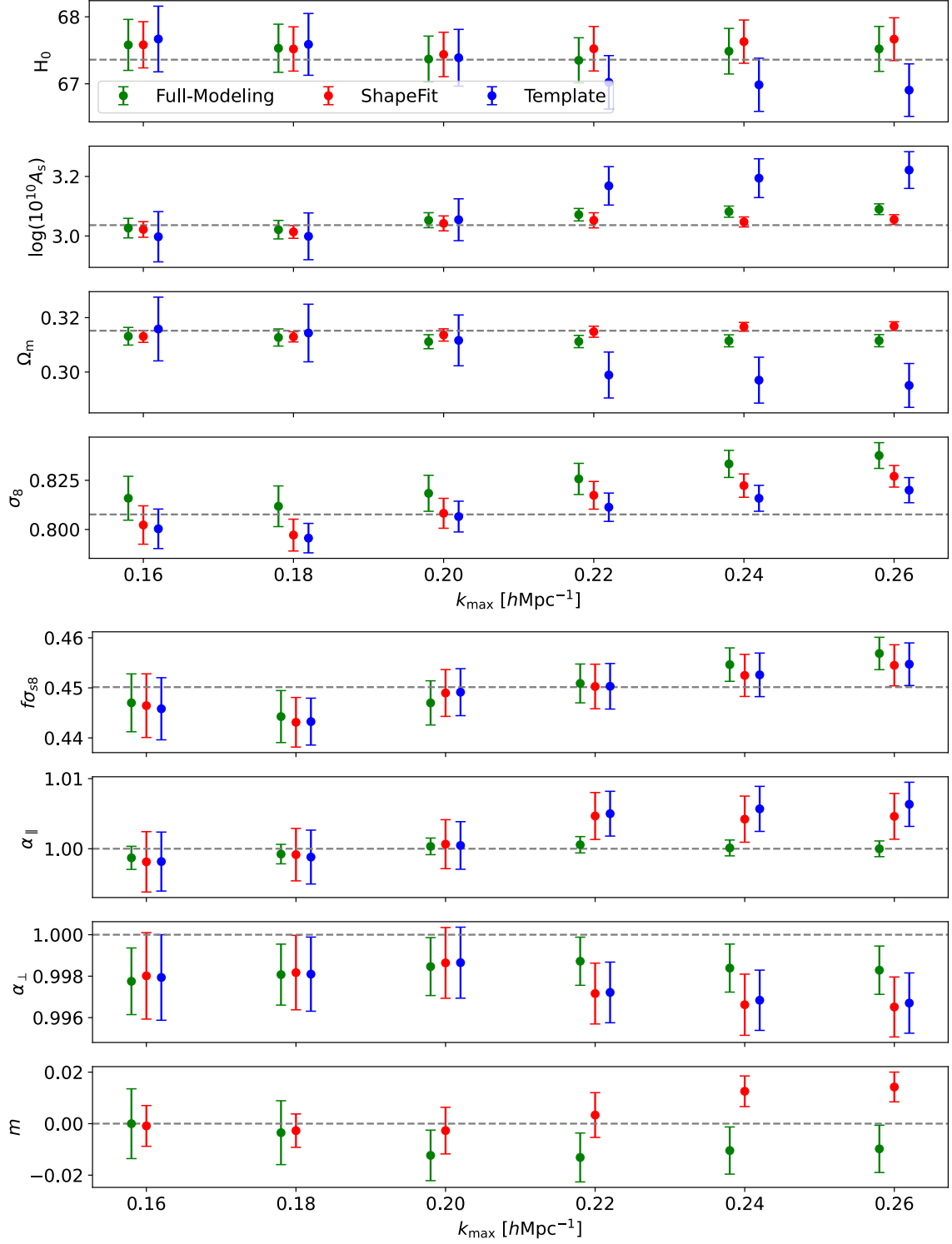
Figure 5. Constraints on ΛCDM and compressed parameters from the three modeling methods with varying $k_{\max}$ between $0.16 \leq k_{\max} \leq 0.26 \ h\mathrm{Mpc}^{-1}$(points slightly offset for clarity). These results are obtained with a covariance appropriate to the 25 box volume, fitting to the LRG cubic mock data.
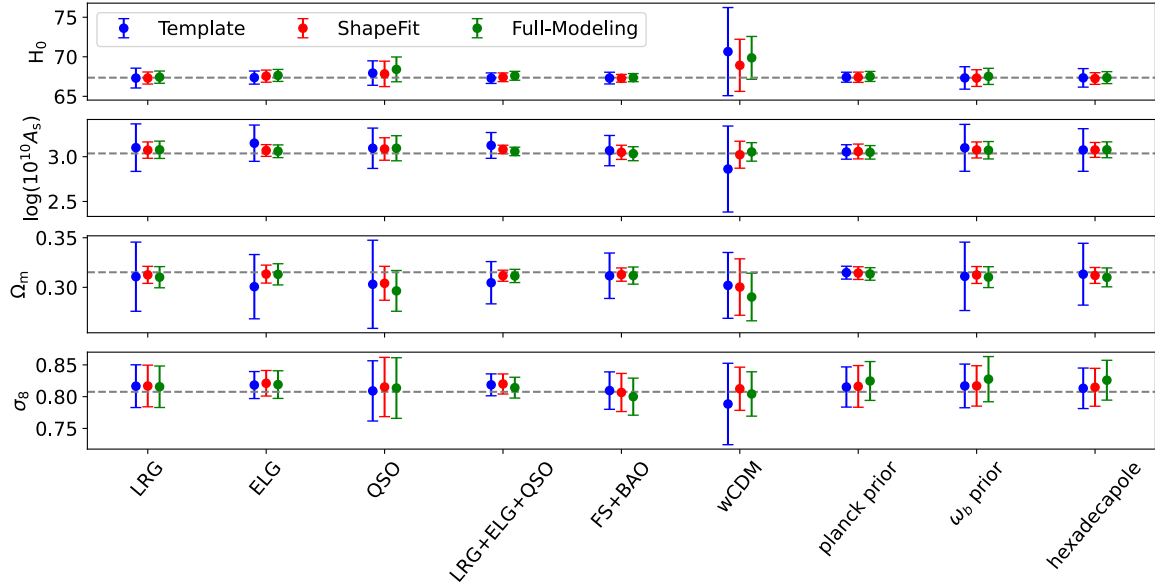
Figure 6. Constraints on ΛCDM parameters for the three modeling methods for a variety of different fit settings and data sets to be discussed in the text. The results are obtained with the covariance for a single box, $8\,(h^{-1}\mathrm{Gpc})^3$, volume and $k_{\max} = 0.2\,h\,\mathrm{Mpc}^{-1}$. In many of the cases Ω_{m} appears slightly below the truth, which is in part due to projection effects. Here "standard" refers to our baseline result on the LRG mocks.

as "summary statistics" of each tracer sample. It is in the cosmological interpretation step that we can either infer ΛCDM parameters from a single sample or from the combination of $(f\sigma_{s8},\alpha_{\parallel,\perp},m)_z$ sets of multiple tracer samples.

In the three panels of figure 7 we show a comparison between results of fitting a single sample versus joint fits of multiple tracers, for the standard template, ShapeFit, and Full-Modeling methods respectively. We observe that in each method, the ELG data is significantly more constraining than the LRG sample, and thus the joint fitting constraints appear to be dominated by the ELG sample. The QSO mocks are the least constraining data set, due to the lower number density of Quasars from which the power spectrum is measured. Therefore the error bars at each Fourier mode are larger than those of the ELG and LRG data, resulting in significantly poorer constraints in the model parameters governing the power spectrum shape, i.e. Ω_{m} and H_0. Meanwhile, the amplitude parameter $\log A_s$ is not as sensitive to the type of tracer and we observe smaller differences in constraint between the tracer types. Overall, the tightest constraints on all parameters are obtained in the joint analysis of LRG+ELG+QSO, but with an almost negligible improvement coming from the inclusion of QSO data.

5.4 Full-shape + BAO reconstruction

In addition to fitting the full-shape power spectra using our model, we can gain extra constraining power through a joint analysis with the reconstructed BAO correlation function. The BAO reconstruction procedure aims to undo some of the damping of the BAO signal due to nonlinear structure growth in order to sharpen its peak, allowing for a better measurement of the cosmological distance-redshift relation via the well-defined drag horizon scale (see
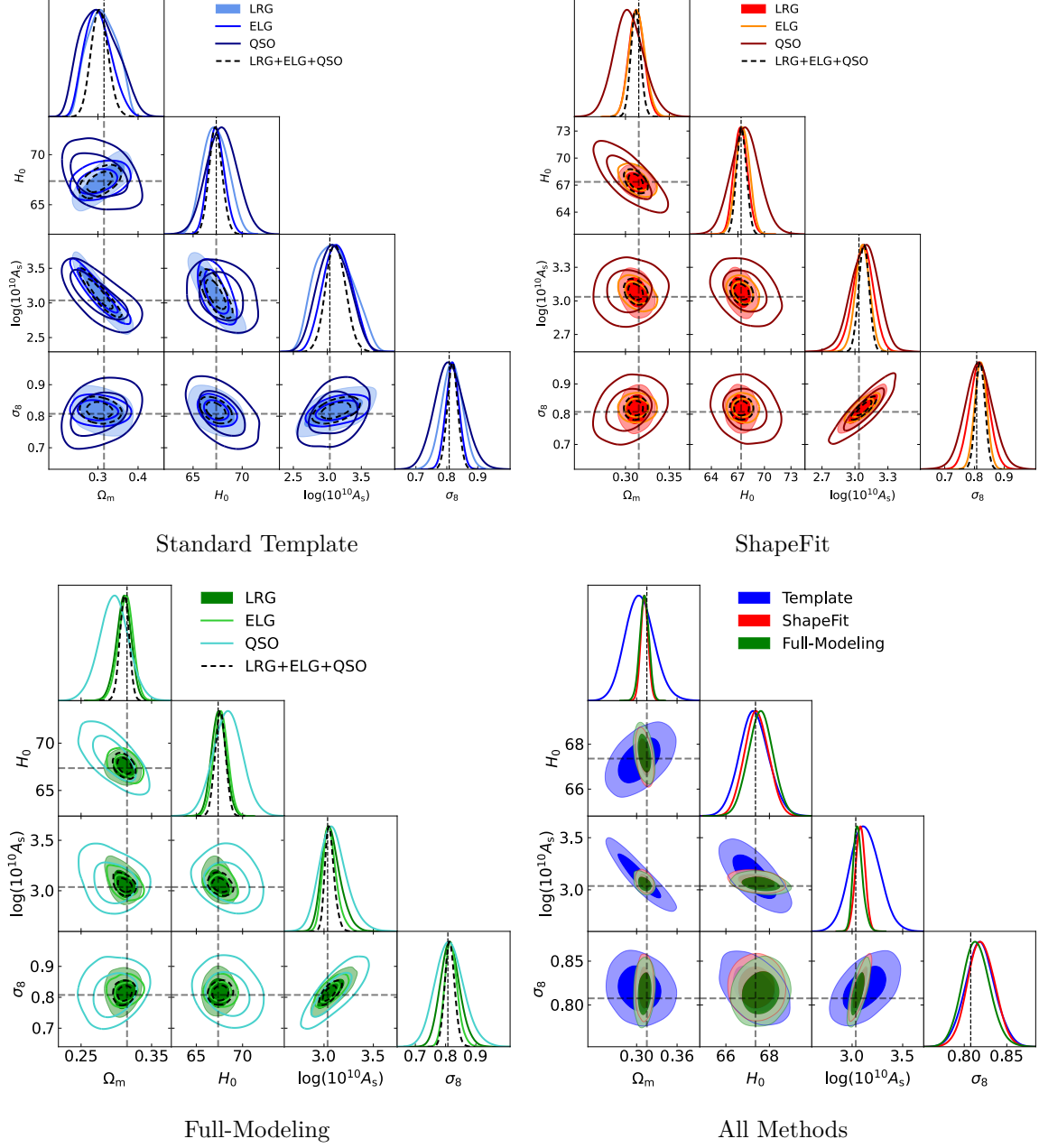
Figure 7. Comparison of cosmological constraints of different tracers (LRG, $z = 0.8$; ELG, $z = 1.1$; QSO, $z = 1.4$) for the three fit methods. Here we only show the results with the covariance of the single-box volume of $(2\,h^{-1}\mathrm{Gpc})^3$ and $k_{\mathrm{max}} = 0.2\,h\,\mathrm{Mpc}^{-1}$.

e.g. refs. [49, 69–73]). This procedure begins by smoothing the observed clustering signal by a Gaussian filter $S(k) = \exp(-(kR_s)^2/2)$, which serves to filter out small-scale modes. Next, we use this smoothed density to estimate the smoothed Zel'dovich displacement, $\boldsymbol{\chi} \approx S(k)\boldsymbol{\Psi}_{\mathrm{Zel}}$, which we subtract from the observed galaxy field as well as from a random matter density field in order to preserve large-scale power. The reconstructed galaxy density field is then $\delta_{\mathrm{rec}} = \delta_{\mathrm{d}} - \delta_{\mathrm{s}}$, with δ_{d} and δ_{s} being the displaced galaxy and shifted random fields, respectively. Moving to redshift space once again amounts to a rotation of the real-space field, $\boldsymbol{\chi}_s = R\boldsymbol{\chi}$ with matrix R defined in section 3. In the literature one commonly encounters two methods for reconstructions in redshift space: **RecSym** [72] and **RecIso** [69, 71]. The first applies the transformation into redshift space equally to both δ_{d} and δ_{s}, whereas the latter method keeps the shifted field in real-space (see ref. [73] for further discussion). For the DESI simulations considered in this work, the **RecSym** procedure is applied to produce the post-reconstruction mock data.

We model the damping of the BAO feature in the reconstructed power spectrum, $P_{\mathrm{rec}} = P_{\mathrm{dd}} + P_{\mathrm{ss}} - 2P_{\mathrm{ds}}$ within the Zel'dovich approximation by splitting the linear theory predictions into the wiggle and no-wiggle components[11] and apply an exponential damping factor[12] to the wiggle part [73]

$$P(k, \mu) = (b + f\mu^2)^2 \left(P^{NW}(k) + e^{-\frac{1}{2}k^2\Sigma^2(\mu)} P_{ab}^W(k) \right), \tag{5.1}$$

where the Σ^2 in the damping factor is the isotropic component of the linear pairwise displacement $A_{ij}^{dd} = \left\langle \Delta_i^{dd} \Delta_j^{dd} \right\rangle$, of the displaced density field at $|\boldsymbol{q}| = r_d$, i.e.

$$\Sigma^2(\mu) = \frac{1}{3}\delta_{ij} A_{ij}(\boldsymbol{q})\Big|_{q=r_d} \tag{5.2}$$

$$= \left[1 + f(2+f)\mu^2 \right] \left[2\tilde{\Sigma}^2(0) - 2\tilde{\Sigma}^2(r_d) \right] \tag{5.3}$$

$$\tilde{\Sigma}^2(q) = \frac{1}{3} \int \frac{dk}{2\pi^2} (1 - \mathcal{S})^2 j_0(kq) P_{\mathrm{lin}}(k). \tag{5.4}$$

Finally, after generating the reconstructed power spectrum, we use a Fourier transform to obtain the reconstructed correlation function. We limit our model to linear bias as it has been found in previous works that the IR damping of the BAO feature dominates over other nonlinear effects such as mode-coupling which are largely cancelled by reconstruction. Following ref. [73] we employ a new method for modeling the broadband that is not degenerate with the BAO signal, which in Fourier space involves using a basis of cubic splines. When fitting the correlation function in configuration space this is equivalent to setting a minimum scale, r_{min}, with the exception of two Hankel transformed basis functions that are included

---

[11]There are numerous methods for performing this split. Here we use the method described in appendix D of ref. [74] that uses a sine transform to identify the BAO feature in real space and subtracts it before transforming back to Fourier space to produce a wigge-free power spectrum.

[12]Previous works studying BAO reconstruction have sometimes derived different damping factors for P_{dd}, P_{ds} and P_{ss}. This results from a 1$^{\mathrm{st}}$ order approximation in LPT, and a more consistent approach has the randoms damped by the same factor. This subtlety is described in detail in ref. [73], as well as in ref. [75] for a slightly different reconstruction scheme. However, we find that the difference between the old and new methods results in negligible effects to the fit posteriors.

in the quadrupole:

$$\mathcal{Q}_{2,n}(r\Delta) = \frac{i^2}{2\pi^2} \int dk\, k^2 W_3\left(\frac{k}{\Delta} - n\right) j_2(kr), \quad n = 0, 1 \tag{5.5}$$

where W_3 is the piecewise cubic spline kernel [76, 77], j_2 is a $\nu = 2$ spherical Bessel function, and we choose $\Delta = 0.06\,h\mathrm{Mpc}^{-1}$ for the separation scale of the splines. We additionally include a template of polynomials in even powers of r for the monopole and quadrupole moments, truncated at quadratic order, to marginalize over contamination by large-scale systematics below some $k_{\min}$. The broadband model in configuration space is thus [73]:

$$\mathcal{B}_0(r) = a_{0,0} + a_{0,1}\left(\frac{rk_{\min}}{2\pi}\right)^2$$
$$\mathcal{B}_2(r) = a_{2,0} + a_{2,1}\left(\frac{rk_{\min}}{2\pi}\right)^2 + \Delta^3(a_{2,2}\mathcal{Q}_{2,0}(r\Delta) + a_{2,3}\mathcal{Q}_{2,1}(r\Delta)) \tag{5.6}$$

where $k_{\min} = 0.02\,h\mathrm{Mpc}^{-1}$ and the parameters $\{a_{0,0}, a_{0,1}, a_{2,0}, a_{2,1}, a_{2,2}, a_{2,3}\}$ can be analytically marginalized over. We use broad Gaussian priors centered at 0 with widths of 5×10^5 for all of these broadband parameters. Finally, we note that one should also include some more flexibility in the damping factor by introducing parameters $\Sigma_{\parallel,\perp}$ in the exponent in eq. (5.1) to marginalize over the effects of nonlinearities. However, we did not find this necessary in the tests presented here, and so the damping factors vary only as f, P_{lin}, and r_{d} change in Full-modeling and likewise with ShapeFit through the $f\sigma_{s8}$ and m parameters.

The joint covariance matrix is computed numerically using the reconstructed correlation function realizations of the `EZmock` simulations. So the joint data vector is now $\mathbf{d} = \{P_0^{\mathrm{pre}}(k), P_2^{\mathrm{pre}}(k), \xi_0^{\mathrm{post}}(r), \xi_2^{\mathrm{post}}(r)\}$ with cross-correlations between $P_\ell^{\mathrm{pre}}(k)$ and $\xi_\ell^{\mathrm{post}}(r)$ accounted for as nonzero off-diagonal elements in the joint covariance matrix. (e.g. see figure 3 of [48])

We show in figure 8 comparisons of the cosmological constraints pre/post BAO reconstruction. We find that for all three modeling methods there is significant improvement in constraints when joint-fitting with the post-recon correlation function, most significantly in H_0 as the cleaner measurement of BAO scale from the sharpened peak allows for better calibration of the distance-redshift relation that constrains Hubble's constant. When comparing all methods we find consistent constraints between ShapeFit and Full-Modeling that are both tighter than those of the standard template.

5.5 Beyond ΛCDM: *w*CDM model

With the expected improvement in cosmological parameter estimation from future galaxy redshift surveys, we hope to place better constraints on parameters not just underlying the standard ΛCDM model, but also departures from it. From the Friedmann equations, the energy density of a specific component of the Universe is related to the scale factor, a, by

$$\rho \propto a^{-3(1+w)} \tag{5.7}$$

where $w = p/\rho$ is the equation of state parameter. One of the simplest extensions to ΛCDM involves allowing the dark energy equation of state to differ from the value of -1 that
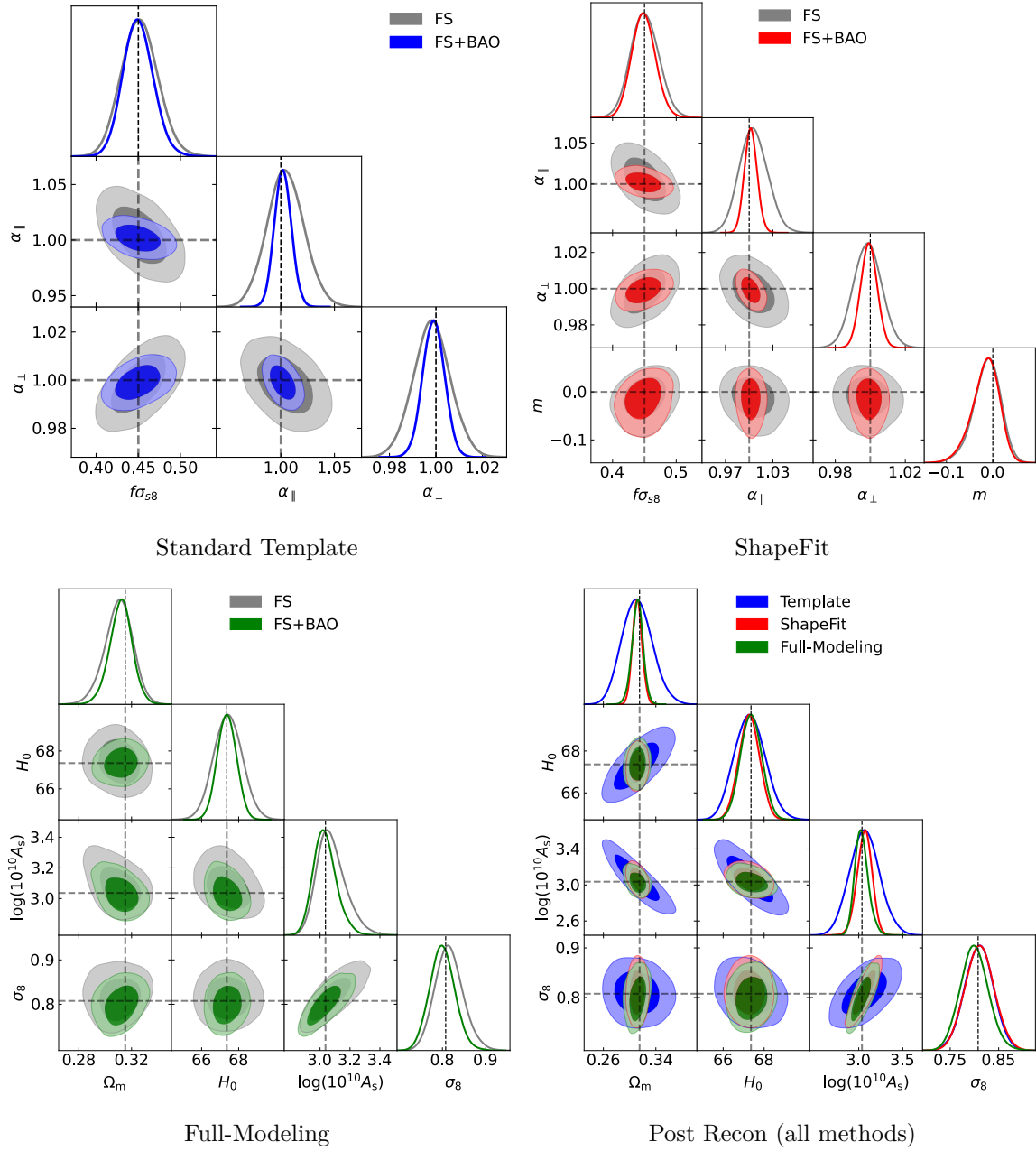
Figure 8. Comparison of cosmological constraints with and without BAO reconstruction for each modeling method. The bottom right panel compares the post-reconstruction constraints of the three methods. For all plots above, we present results using the covariance appropriate to the single-box volume. In the legends, "FS" refers to the pre-reconstruction full-shape power spectrum data, and "BAO" refers to the BAO signal in the post-reconstruction correlation function.
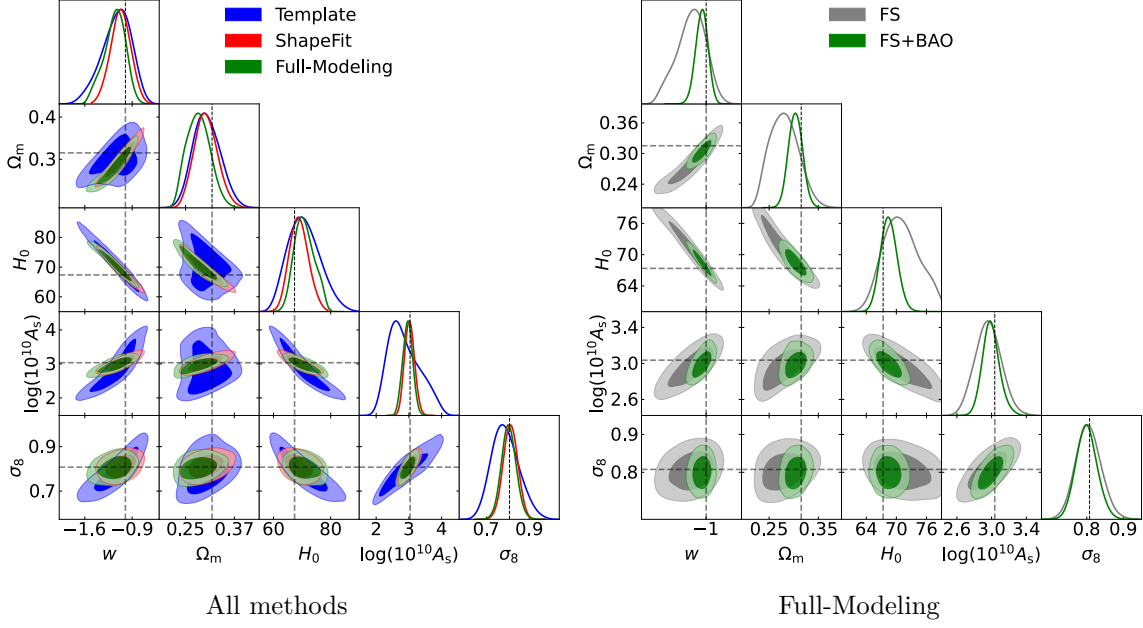
Figure 9. On the left we show a comparison of cosmological constraints on the wCDM parameters for the three modeling methods. On the right we show wCDM Full-Modeling fits with and without the inclusion of post-reconstruction BAO data. In both plots we show the results with the single-box $(2\,h^{-1}\mathrm{Gpc})^3$ volume. In the legends, "FS" refers to the pre-reconstruction full-shape power spectrum data, and "BAO" refers to the BAO signal in the post-reconstruction correlation function.

corresponds to a cosmological constant (Λ) as the energy density is constant in that case. On the other hand, "quintessence" models have $w \neq -1$ such that dark energy is a dynamic quantity in the Universe.[13]

Figure 9 shows in the left panel the constraints on wCDM parameters obtained from each of the three modeling methods, for the covariance of the single-box volume. Since the Abacus cosmology assumes a cosmological constant for dark energy, the expected value is -1. We find that the ShapeFit and Full-Modeling methods both give constraints on w that are within 1σ of the expected equation of state. Meanwhile the parameters in the template method are very poorly constrained when w is varied. When changing the properties of dark energy away from the cosmological constant the universe's expansion history and geometry are significantly altered, thus affecting the $\alpha_{\parallel,\perp}$ parameters and $f\sigma_{s8}$. This results in the observed degeneracies between w and the other parameters (which also determine $f\sigma_{s8}$, and $\alpha_{\parallel,\perp}$). If those three parameters are the only information we have from the data, as is the case in the template fit, then this results in very poor constraints. However, moving far along those degeneracy

---

[13]If dark energy is described by a scalar field, ϕ, with a canonical kinetic term then the equation of state can be interpreted in terms of kinetic and potential energies via,

$$w = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}.$$

Under this assumption the equation of state is usually expected to lie between $-1 < w < 1$, with values $w < -1/3$ leading to cosmic acceleration. However, more exotic models exist that do allow for negative kinetic energies.

directions also significantly affects the shape of the power spectrum, which the ShapeFit and Full-Modeling methods *are* sensitive to. Therefore these two methods do not suffer from the degeneracies as much as the template fit. Comparing ShapeFit to Full-Modeling, we find that the constraints on parameters from the ShapeFit method are a bit wider than in Full-Modeling. This is likely because all of the shape information is contained in a single parameter, which then needs to be interpreted as constraints on three different cosmological parameters ($w, \omega_{\rm m}$, and H_0), as these all control the shape of the power spectrum. Thus, a poorer measurement of m results in more sensitivity to the degeneracies in shape that the template fit also suffered from. Finally, we also note that projection effects (see appendix B) in Full-Modeling cause close-to 1σ offsets in the $w, \Omega_{\rm m}$, and H_0 parameters. While these shifts are not huge for this dataset, we also are interested into what extent including more data can mitigate projection effects. We show in the right panel of figure 9 a comparison of Full-Modeling fits with and without the inclusion of reconstructed BAO data. We find that including BAO results in noticeable improvements in the constraints by shifting the posteriors closer to the truth. These projection effects are not as significant in the ShapeFit method, which suggests that the extra information that Full-modeling obtains w.r.t. ShapeFit may come from regions of the power spectrum that are degenerate with counterterm and/or stochastic parameters. A similar effect was observed and reported in ref. [52] when comparing $f\sigma_{s8}$ constraints between Full-Modeling and standard template methods in BOSS data.

5.6 Priors from CMB

The 'standard' template method was conceived at a time when the data from galaxy redshift surveys was not constraining enough on early-universe physics to be competitive with constraints from probes such as Planck that modeled CMB anisotropies. In particular, data from CMB anisotropies tightly constrain the ΛCDM parameters that determine the shape of the power spectrum [78], and this shape is left unaltered by late-time physics such as dark energy or spatial curvature. These constraints are tighter than those from the galaxy survey themselves. In such a scenario, the primary degrees of freedom to be constrained by galaxy surveys are late time growth and the late-time distance-redshift relation. The template method was intended to be used in conjunction with the other probes, such that most of the information on $P(k)$ shape came from strong priors using results from e.g. Planck. To demonstrate this, we repeat the cosmological inference of the template results, but including an additional likelihood derived from the Planck 2018 results [79]. We do this by taking the chains from the baseline model of the Planck Legacy Archive, "base plikHM TT lowl lowE", and compute the covariance matrix, $C_{\rm pl}$, from the ($\omega_{\rm b}$, $\omega_{\rm cdm}$) samples. We do not apply a prior on A_s or H_0 as we are interested in how information from galaxy clustering constrains the late-time growth compared to Planck. When we sample in these ΛCDM parameters we now include the additional likelihood

$$\mathcal{L}_{\rm pl} \propto \exp\left\{-\frac{1}{2}\boldsymbol{\Delta\Theta}^T C_{\rm pl}^{-1}\boldsymbol{\Delta\Theta}\right\}, \tag{5.8}$$

where $\boldsymbol{\Delta\Theta}$ is the difference between the sampled ($\omega_{\rm b}$, $\omega_{\rm cdm}$) and the values in the Abacus cosmology. Because we are including the CMB prior on ω_b, we remove the BBN prior that we usually use in our standard analyses. We show these results, comparing the template,
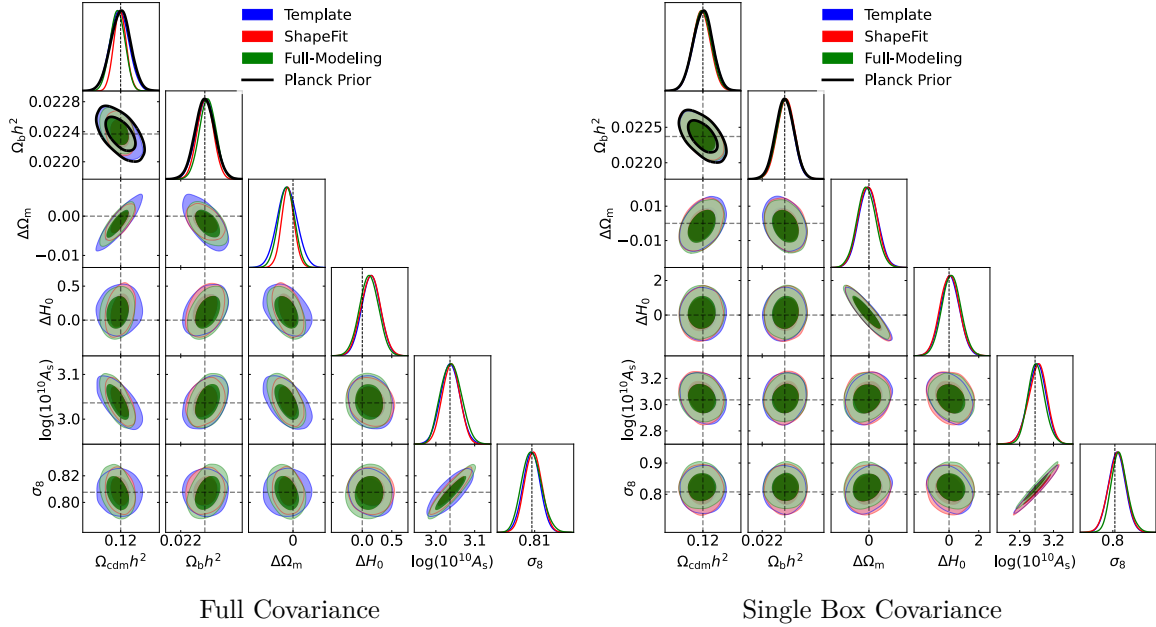
Figure 10. Comparison of ΛCDM constraints from the template fit(blue), Full-Modeling (green), and ShapeFit (red) with Planck priors, on the Abacus LRG ($z = 0.8$) mock data. On the left we show the fits with the covariance for the full ($25 \times (2\,h^{-1}\mathrm{Gpc})^3$) volume while the right figure shows the results with the single-box $(2\,h^{-1}\mathrm{Gpc})^3$ volume. In the $\Omega_{\mathrm{cdm}}h^2 - \Omega_{\mathrm{b}}h^2$ panels we include black lines showing the Planck prior.

Shapefit, and Full-Modeling methods with Planck priors, in figure 10, using the LRG ($z = 0.8$) mock data within the standard ΛCDM model. We see that the inclusion of Planck priors significantly tightens the constraints on Ω_{m}. Despite us not applying any prior on H_0 and $\log A_s$, we still observe a shift to the truth and tightening in those parameter constraints for all three methods, with the $\log A_s$ posterior slightly narrower for the Full-Modeling approach. Overall, all three methods agree very closely in all of the parameters when including these priors, suggesting that the difference in constraining power of these methods is almost entirely due to shape information (which is better determined by the CMB than the galaxy survey).

5.7 Varying n_s

For previous fullshape analyses from spectroscopic surveys, it was common/necessary to fix (or impose tight priors) on several of the ΛCDM parameters such as ω_b, M_ν, and n_s, using information from the CMB and BBN. With the increasing constraining power of DESI and future surveys it is of interest to see how much we can untangle fullshape analyses from other probes. While a tight prior on ω_b (see appendix D) is still necessary, the improved constraining power of DESI may allow us to free n_s and/or M_ν.[14] To investigate the impact of uncertainty in n_s on our analysis given the statistical uncertainties in Y1, we chose mock data from one of the DESI Y1 redshift bins (LRG; $0.4 < z < 0.6$) with an appropriate analytic covariance. We compare constraints on ΛCDM parameters with various prior choices on n_s, including a uniform prior, Gaussian with widths of $10\times$ and $5\times$ Planck 2018 constraints

---

[14]In this paper we only perform tests with n_{s} free and refer readers to ref. [44] for a discussion on varying M_ν.
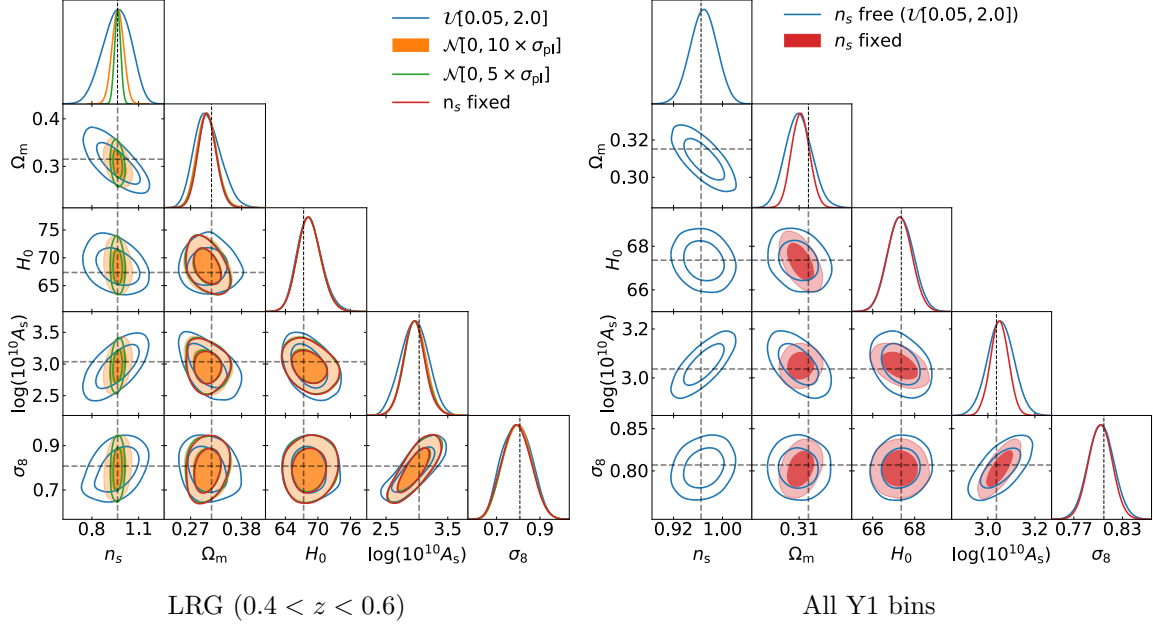
Figure 11. The left panel shows Full-Modeling fits to the mean of Cutsky mocks in the DESI Y1, LRG $0.4 < z < 0.6$ redshift bin, with different priors on n_s. We use the notation $\mathcal{U}[\text{min}, \text{max}]$ and $\mathcal{N}[\mu, \sigma]$ to denote uniform and Gaussian priors, respectively. In the right panel we compare Full-Modeling n_s fixed versus free using the synthetic mocks created with velocileptors simulating the full Y1 footprint: BGS $(0.1 < z < 0.4)$, LRG $(0.4 < z < 0.6)$, LRG $(0.6 < z < 0.8)$, LRG $(0.8 < z < 1.1)$, ELG $(0.8 < z < 1.1)$, ELG $(1.1 < z < 1.6)$, and QSO $(0.8 < z < 2.1)$. We show constraints with uniform priors on n_s in the free case.

$(\sigma_{ns} = 0.004)$ [79], and with n_s fixed. These results are shown in the left panel of figure 11 for the Full-Modeling method. We find that for both the $10\times$ and $5\times$ priors on n_s the constraints on ω_{cdm}, H_0, and $\log A_s$ are identical to those when n_s is fixed, suggesting that the Full-Modeling constraints on ΛCDM parameters are robust even if the n_s constraints from the CMB are systematically off by 10σ. In order to see how well n_s can be constrained completely independently from Planck we additionally fit to noiseless synthetic mock data vectors simulating all seven DESI Y1 redshift bins: BGS $(0.1 < z < 0.4)$, LRG $(0.4 < z < 0.6)$, LRG $(0.6 < z < 0.8)$, LRG $(0.8 < z < 1.1)$, ELG $(0.8 < z < 1.1)$, ELG $(1.1 < z < 1.6)$, and QSO $(0.8 < z < 2.1)$ using the appropriate Y1 analytic covariance for each redshift bin. We compare the case with uniform priors on n_s to the case with n_s fixed. These results are shown in the right panel of figure 11. We find that despite the slight degradation in Ω_{m} constraint with the flat prior on n_s, we are able to measure n_s to a 3% precision.

5.8 Comparison of LPT and EPT

In addition to the LPT model that we primarily focus on in this paper, `velocileptors` also has an Eulerian perturbation theory module. The EPT kernels are constructed from the Lagrangian kernels while setting the IR resummation scale, k_{IR}, to zero. The Eulerian and Lagrangian theories differ in their treatment of cold dark matter, the first describing dark matter as a perfect pressureless fluid, and the latter describing it as collisionless particles.

The overdensities derived from both theories agree order-by-order except when particle trajectories cross. The EPT model in `velocileptors` employs the galaxy bias scheme described in ref. [80]. The mapping between the Lagrangian and Eulerian bias bases can be achieved within `velocileptors` via the transformations [37]:

$$
\begin{aligned}
b_1^E &= 1 + b_1^L \\
b_2^E &= b_2^L + \frac{8}{21}b_1^L\,, \quad b_s^E = b_s^L - \frac{2}{7}b_1^L \\
b_3^E &= 3b_3^L + b_1^L.
\end{aligned}
\tag{5.9}
$$

Lastly, the IR resummation in EPT is performed by splitting the wiggle and no-wiggle parts of the power spectrum, using the same method as is employed in modeling the poste-reconstruction BAO correlation function (section 5.4) and applying a damping factor to the wiggle component. We refer readers to ref. [37] for full details of the Eulerian model and how it compares to LPT. We show in figure 12 a comparison of Full-Modeling constraints when fitting the LRG cubic mocks using LPT and EPT. We see that the constraints agree to within fractions of a σ. A more detailed comparison between the two models, including fits to the ELG and QSO mocks for ShapeFit and Full-Modeling, is presented in ref. [47] along with comparisons to other EFT models on the market.

5.9 Varying f and σ_8 separately

The "standard" method of compression involves varying f while keeping σ_{s8} fixed to the fiducial value σ_8^{ref}, and then reporting the product as $f^{\mathrm{true}}\sigma_{s8}^{\mathrm{true}}$. In principle, one should be able to vary f and σ_{s8} independently and present the result as $f^{\mathrm{true}}\sigma_8^{\mathrm{true}}$. This is because the degeneracy between f and σ_{s8} is broken in the 1-loop terms of the power spectrum. In order to test the ability to constrain σ_{s8}, we run a fit in which $\sigma_{s8}(z=0)$ is a free parameter in addition to $f(z)$ and the other compressed parameters. We vary $\sigma_{s8}(z=0)$ by re-scaling the linear power spectrum by:

$$
P'_{\mathrm{lin}}(k) = \left(\frac{\sigma_{s8}}{\sigma_8^{\mathrm{fid}}}\right)^2 \times P_{\mathrm{lin}}(k),
\tag{5.10}
$$

where $\sigma_8^{\mathrm{fid}} = 0.8076$ for the Abacus fiducial cosmology. The reported $f\sigma_{s8}$ is then $f\sigma_{s8} = f(z)\sigma_{s8}D(z)$ where the growth factor $D(z)$ is computed from the fiducial value of $\Omega_{\mathrm{m}} = 0.315$. We show these results in figure 13. We observe that even though $f\sigma_{s8}$ agrees with that obtained from the standard method, the σ_{s8} constraint of 0.570 ± 0.087 is significantly below the true value of 0.8076. This implies a growth rate $f(z) \sim [\Omega_{\mathrm{m}}(z)]^{0.55} > 1$ which is unphysical. While it is unfortunate that the 1-loop corrections to the power spectrum can not sufficiently constrain f and σ_{s8} independently, we reiterate that our constraint on $f\sigma_{s8}$ remains robust. We also note a slight degeneracy between σ_{s8} and m. While m is designed to change the shape of the power spectrum, σ_{s8} is an integrated quantity that is also mildly affected by changes in the shape.
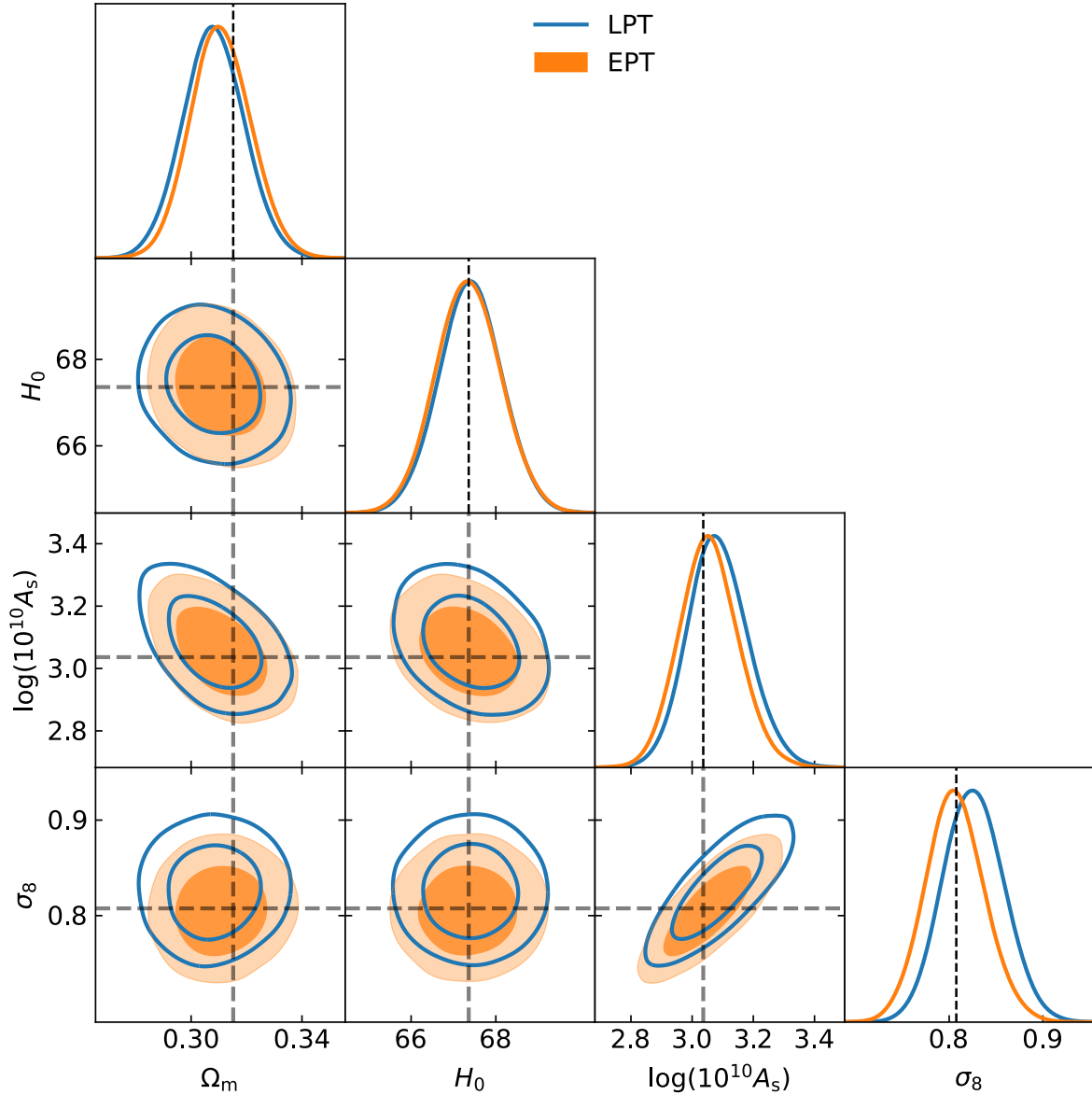
Figure 12. Full-Modeling fits to the mean of LRG Cubic mocks using the LPT and EPT models within `velocileptors`.

6 Conclusion

Observations are probing the Universe and its evolution with unprecedented precision, allowing for significant improvements in measurements of fundamental parameters. The increased constraining power of these data also increases the sensitivity of our results to systematic effects present in models and analysis methods. The largest galaxy redshift survey to date, the Dark Energy Spectroscopic Instrument (DESI), is currently under way with its first year of Fullshape data being unblinded in the spring of 2024. To prepare for unblinding we must have a detailed understanding of the sources of systematic and theoretical error when fitting observations, the flexibility and limitations of our models, and the performance of different analysis methods. In this paper we presented tests of these effects using the public
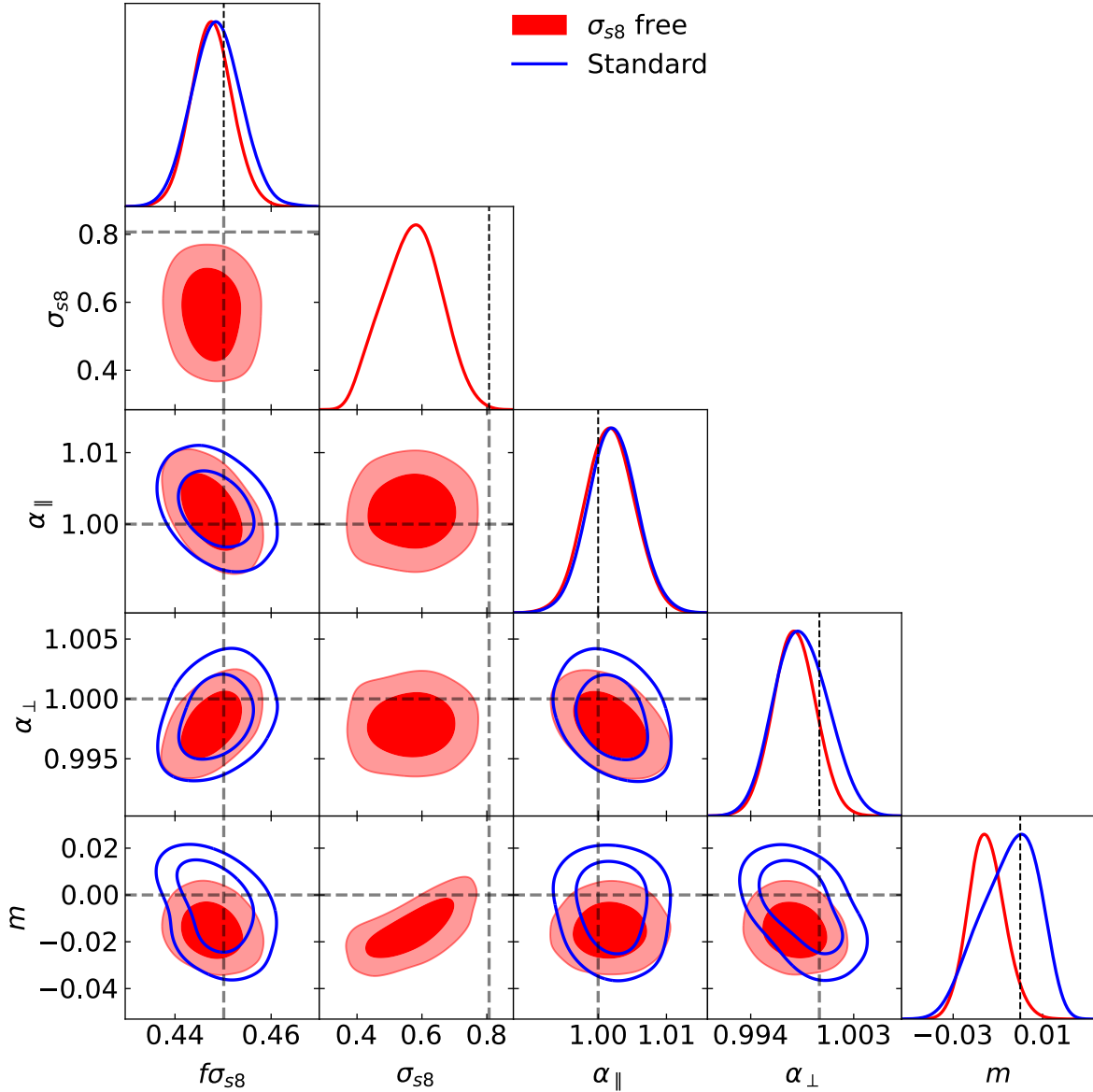
Figure 13. ShapeFit constraints, where in the red data we allow $\sigma_{s8}(z=0)$ to vary independently of $f(z)$. This test is performed using the covariance for the 25 box volume of 200 $(h^{-1}\mathrm{Gpc})^3$.

effective-perturbation-theory code `velocileptors`, fitting data from the AbacusSummit suite of simulations. Our focus will be on cosmological constraints using the Lagrangian Perturbation Theory (LPT) module in `velocileptors`, though we also explore fits using its Eulerian Perturbation Theory (EPT) counterpart. In particular, we fit LRG, ELG, and QSO mock data at effective redshifts of $z = 0.8, 1.1, 1.4$ respectively, consisting of clustering measurements from 25 cubic boxes of 8 $(h^{-1}\mathrm{Gpc})^3$ each for a total volume of 200 $(h^{-1}\mathrm{Gpc})^3$ for each tracer type. Companion papers to this one, using other effective perturbation theory codes `Folps`ν and `PyBird`, are scheduled to appear concurrently (refs. [44, 45], including in addition a comparison paper (ref. [47]) showing that all three effective-theory pipelines and models behave very similarly when the underlying assumptions and settings are consistent.

In this paper we discussed three modeling methods: (1) the standard Template fit, the default method used in previous BOSS and eBOSS analyses, that compresses observed multipoles into three summary statistics, ($f\sigma_{s8}$,$\alpha_\parallel$,$\alpha_\perp$) while keeping the linear power spectrum fixed; (2) the ShapeFit method which introduces an additional compressed parameter m to the standard Template that modulates the shape of the linear template power spectrum which depends on early universe physics; and (3) the Full-Modeling method which directly samples in the parameter space of a cosmological model in order to fit the data. The first two methods are model-agnostic and so the compression only needs to be performed once, after which the obtained summary statistics can be mapped to any cosmological model (ΛCDM or extensions) of ones choosing. Despite the Full-Modeling method technically requiring a Boltzmann code to compute the linear power spectrum at every step of an MCMC, the use of Taylor series expansion emulators make the difference in computational cost/time negligible when compared to the compressed analyses.

We showed throughout the paper that the increased information from the shape of the linear power spectrum results in significant improvements in cosmological constraints in ShapeFit when compared to the standard Template analysis, when CMB data are not included. Compared to the Full-Modeling approach, ShapeFit provides consistent results on ΛCDM (and wCDM) parameters with minimal loss in constraining power. In varying the upper bound of the fitting range, we found that the models give unbiased constraints for scale cuts up to $k_{\max} \leq 0.2\,h\mathrm{Mpc}^{-1}$. When including priors from Planck in order to constrain early universe information, all three methods give consistent results. Since the upcoming data will include tracers from different redshifts, we tested the ability of our pipelines in fitting simultaneously the tracers from three redshift bins, finding the joint analysis to improve the constraints without any noticeable systematic effects.

Because one of the most powerful sources of cosmological information in LSS that DESI can detect is the Baryon Acoustic Oscillation (BAO) signal, whose well-defined scale can be used as a standard ruler to constrain the distance-redshift relation, we combined our fullshape analyses with post-reconstruction BAO correlation function, finding significant improvements in constraints for each modeling method. Finally, we also show how each method performs when extending the parameter space beyond the standard ΛCDM model by varying the dark energy equation of state parameter w. The ShapeFit and Full-Modeling methods are both able to obtain consistent and unbiased constraints within the wCDM model, whereas the standard template suffers greatly from degeneracies that can not be broken without shape information.

In addition to the `velocileptors`LPT model, the pipeline also has a module based on Eulerian perturbation theory (EPT). We show that these two theoretical frameworks provide consistent constraints, in agreement with the more extensive comparisons along with other PT pipelines, `FOLPS`ν and `PyBird`, presented in ref. [47].

We conclude by summarizing the optimal setup for `velocileptors` for DESI Y1 fullshape analyses. The scaling of the biases with σ_8 appears to be a more natural choice of parameterization that is closer to the constraints from the data and can ameliorate shifts to lower σ_8 in the posteriors when the data is not sufficiently constraining. We recommend against the use of the partial Jeffrey's prior in attempts to reduce projection effects, due to it being a highly informative prior in the cosmological parameters. Our counterterm parameterization

that scales relative to linear theory allows for a more intuitive choice of priors on the α_n parameters as "fractional corrections to linear theory". When fitting the hexadecapole we strongly suggest restricting the k-range in P_4 to a $k_{4,\text{max}} \sim 0.1\ h\text{Mpc}^{-1}$ as this minimizes the model's sensitivity to higher orders in perturbation theory and non-linear effects such as Fingers of God. For the monopole and quadrupole a scale cut of $k_{\text{max}} = 0.20\ h\text{Mpc}^{-1}$ has been found to perform well. Finally, we also suggest the use of physically motivated Gaussian priors on the stochastic parameters that can be justified based on the characteristic physical scales in the system (as captured, for example, in the halo model).

7 Data availability

Data from the plots in this paper are available on Zenodo as part of DESI's Data Management Plan (DOI: 10.5281/zenodo.10951714). The data used in this analysis will be made public along the Data Release 1 (details in https://data.desi.lbl.gov/doc/releases/).

A Analytic marginalization

We can substantially speed up our MCMC fits by analytically marginalizing over the linear nuisance parameters in our model, i.e. the parameters of the stochastic and counterterm contributions (α_0, α_2, α_4, SN_0, SN_2, SN_4). By reducing the number of sampled parameters our chains are able to converge in under 10 minutes instead of an hour or two. The procedure for marginalizing over the linear parameters b_i involves splitting the theoretical prediction, into the piece dependent on the nonlinear parameters $\boldsymbol{a}$ that we sample in and the "template" piece that is multiplied by the linear parameters: $\boldsymbol{\Psi} = \boldsymbol{\Psi}_0(\boldsymbol{a}) + \sum_i \theta_i \boldsymbol{\Psi}_{t,i}$. The likelihood distribution marginalized over the linear nuisance parameters is given by [81, 82]

$$P(\boldsymbol{\Psi}_d | \boldsymbol{\Psi}_0, \boldsymbol{\Psi}_t, \sigma_\theta) = \int d\boldsymbol{\theta}\ \mathcal{L}(\boldsymbol{\Psi}_d | \boldsymbol{\Psi}_0, \boldsymbol{\Psi}_t, \boldsymbol{\theta}) P(\boldsymbol{\theta}), \tag{A.1}$$

where $\boldsymbol{\Psi}_d$ is the data and $P(\boldsymbol{\theta})$ denotes the priors on parameters θ_i, which we choose to be Gaussian (centered at zero) with widths $\sigma_{\theta,i}$:

$$P(\theta_i | \sigma_{\theta,i}) = \frac{1}{\sqrt{2\pi \sigma_{\theta,i}^2}} \exp\left(-\frac{\theta_i^2}{2\sigma_{\theta,i}^2} \right) \tag{A.2}$$

The model likelihood in the integrand is

$$\mathcal{L}(\boldsymbol{\Psi}_d | \boldsymbol{\Psi}_0, \boldsymbol{\Psi}_t, \boldsymbol{\theta}) = (2\pi)^{-n/2} \left| \mathcal{C}^{-1} \right|$$
$$\times e^{-\frac{1}{2}\left[\boldsymbol{\Psi}_d - \left(\boldsymbol{\Psi}_0 + \sum_i \theta_i \boldsymbol{\Psi}_{t,i} \right) \right]^{\mathrm{T}} \mathcal{C}^{-1} \left[\boldsymbol{\Psi}_d - \left(\boldsymbol{\Psi}_0 + \sum_i \theta_i \boldsymbol{\Psi}_{t,i} \right) \right]}. \tag{A.3}$$

Defining $\boldsymbol{\Delta} = \boldsymbol{\Psi}_d - \boldsymbol{\Psi}_0$ and $\log \mathcal{L}_0 = -\frac{1}{2} \boldsymbol{\Delta}^{\mathrm{T}} \mathcal{C}^{-1} \boldsymbol{\Delta}$ we get

$$P(\boldsymbol{\Psi}_d | \boldsymbol{\Psi}_0, \boldsymbol{\Psi}_t, \sigma_\theta) \propto \mathcal{L}_0 \int d\boldsymbol{\theta}\ e^{-\frac{1}{2} \sum_{i,j} \theta_i \theta_j \left(\boldsymbol{\Psi}_{t,i}^{\mathrm{T}} \mathcal{C}^{-1} \boldsymbol{\Psi}_{t,j} + \frac{1}{\sigma_i \sigma_j} \delta_{ij} \right) + \sum_i \boldsymbol{\Delta}^{\mathrm{T}} \mathcal{C}^{-1} \theta_i \boldsymbol{\Psi}_{t,i}}$$
$$= \mathcal{L}_0 \int d\boldsymbol{\theta}\ \exp\left[-\frac{1}{2} \left(\boldsymbol{\theta}^{\mathrm{T}} L \boldsymbol{\theta} - V^{\mathrm{T}} L^{-1} V \right) \right]$$
$$= \frac{(2\pi)^{n/2}}{\sqrt{|L|}} \mathcal{L}_0 e^{\frac{1}{2} V^{\mathrm{T}} L^{-1} V}, \tag{A.4}$$

where we completed the square in the second line and defined the matrices $L_{ij} = \boldsymbol{\Psi}_{t,i}^{\mathrm{T}} \mathcal{C}^{-1} \boldsymbol{\Psi}_{t,j} + \delta_{ij}/(\sigma_i \sigma_j)$ and $V_i = \boldsymbol{\Psi}_{t,i}^{\mathrm{T}} \mathcal{C}^{-1} \boldsymbol{\Delta}$ before taking the multivariate Gaussian integral. So then the log-likelihood consists of the four terms

$$\log P = \log \mathcal{L}_0 + \frac{1}{2} V^{\mathrm{T}} L^{-1} V - \frac{1}{2} \log |L| + \frac{n}{2} \log(2\pi). \tag{A.5}$$

Despite analytically marginalizing over the linear parameters, we can always recover their distribution using the chain containing non-linear parameters. At each step of the chain, the nonlinear parameters are fixed and the likelihood is a Gaussian function of the linear parameters with known mean and variance, i.e. for step n in the MCMC, the likelihood depends on linear parameter θ_i like:

$$\log \mathcal{L}_{n,i} = (\theta_i - \bar{\theta}_i)^{\mathrm{T}} \mathcal{N}^{-1} (\theta_i - \bar{\theta}_i) + \mathrm{const} \tag{A.6}$$
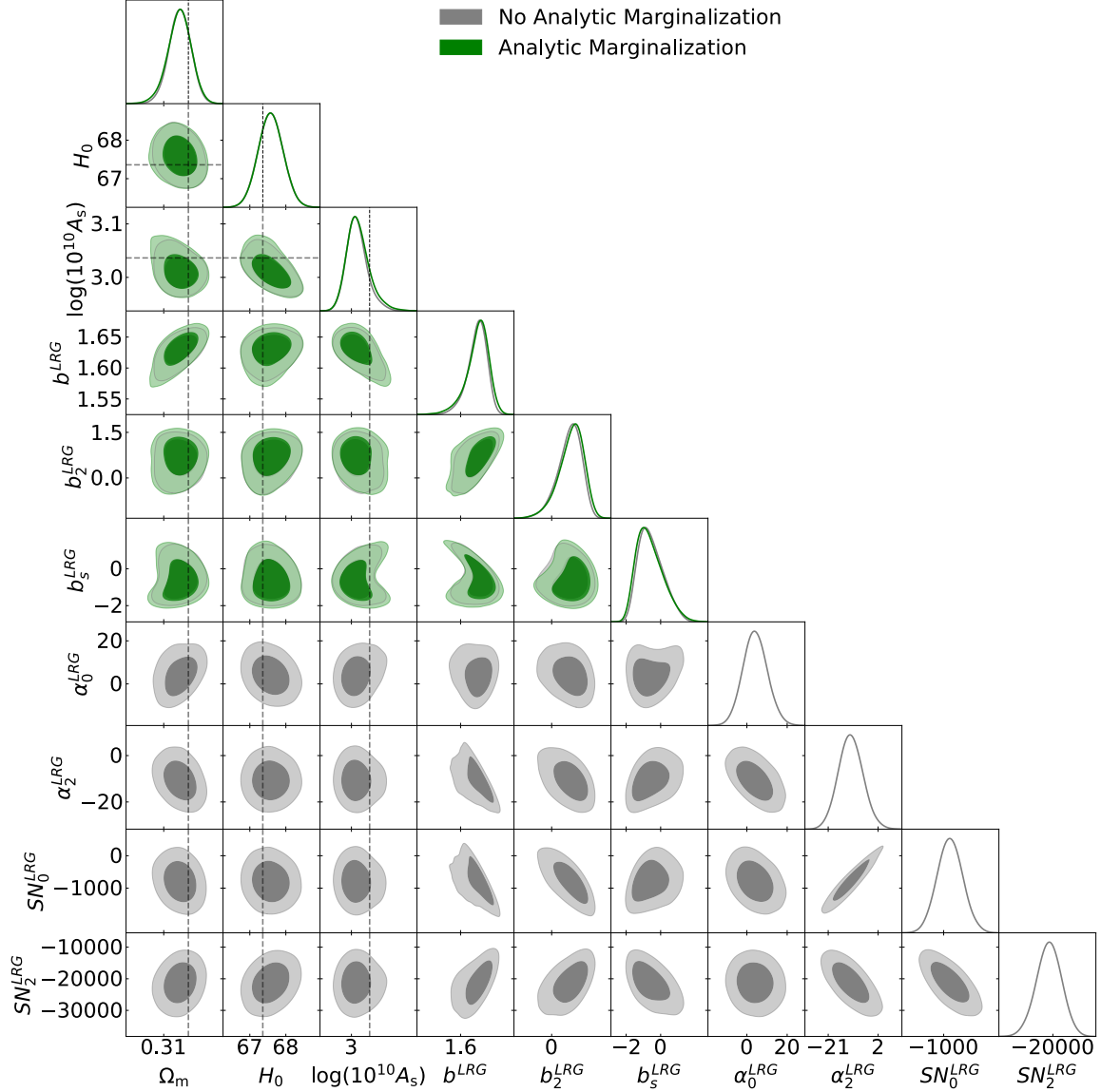
Figure 14. Comparison of Full-Modeling constraints with and without analytic marginalization of linear parameters: α_0, α_2, SN_0, SN_2. Here we show results for the full $25 \times (2\,h^{-1}\mathrm{Gpc})^3$ volume to show that our analytical marginalization method is robust even at very large volumes.

with variance $\mathcal{N}$ and the mean $\bar{\theta}_i$ determined by the (fixed) non-linear parameters. Reconstructing the distribution of parameter θ_i simply amounts to averaging over all of these Gaussians. This allows us to still be able to e.g. check the effects of our priors or to identify any degeneracies between linear parameters and others in the model that could be driving projection effects. We show in figure 14 a comparison of constraints from the Full Modeling method with and without analytic marginalization of the linear parameters. For the parameters that are being sampled in both cases, we find consistent behavior in the contours as expected. In order to make sure that the analytic marginalization is also correctly handling the parameters that we marginalize over, we maximize the first two terms in (A.5) (the latter

	Params	FM Standard (σ)	FM Analytic Marg (σ)
Non-linear	H_0	67.67 (0.35)	67.63 (0.34)
	Ω_{m}	0.3139 (0.0023)	0.3143 $\left(^{+0.0026}_{-0.0023}\right)$
	$\log(10^{10} A_{\mathrm{s}})$	2.998 $\left(^{+0.017}_{-0.023}\right)$	3.001 $\left(^{+0.017}_{-0.026}\right)$
	$b\sigma_8$	1.642 $\left(^{+0.019}_{-0.013}\right)$	1.644 $\left(^{+0.022}_{-0.013}\right)$
	b_2	0.8982 $\left(^{+0.49}_{-0.32}\right)$	0.8705 $\left(^{+0.51}_{-0.33}\right)$
	b_s	-0.7607 $\left(^{+0.55}_{-0.87}\right)$	-0.8512 $\left(^{+0.55}_{-0.95}\right)$
Linear	α_0	0.6987 (6.1)	2.468
	α_2	-11.69 (5.7)	-13.08
	SN_0	-890.3 (420)	-962.4
	SN_2	$-1.919\mathrm{e}4$ (4300)	$-1.911\mathrm{e}4$

Table 3. Comparison of Full-Modeling best-fit parameters with and without analytic marginalization. Uncertainties of the posterior distributions are given in parentheses for all sampled parameters.

terms describe the volume/width of the likelihood surface). This gives us the best-fitting values for the nonlinear parameters. From the maximized posterior, the corresponding best-fit points of the analytically marginalized parameters can then be directly calculated:

$$\theta_j^{\mathrm{bf}} = \sum_i V_i L_{ij}^{-1}. \tag{A.7}$$

Once we have found the best-fitting nonlinear parameters and by extension $\boldsymbol{\Psi}^{\mathrm{bf}} = \boldsymbol{\Psi}_0^{\mathrm{bf}} + \sum_i \theta_i^{\mathrm{bf}} \boldsymbol{\Psi}_{t,i}^{\mathrm{bf}}$, the maximum log-likelihood is just:

$$\log P^{\mathrm{max}} = -\frac{1}{2}[\boldsymbol{\Psi}^{\mathrm{bf}}]^{\mathrm{T}} \mathcal{C}^{-1} \boldsymbol{\Psi}^{\mathrm{bf}} + \log|\mathcal{C}^{-1}| - \frac{n}{2}\log(2\pi). \tag{A.8}$$

In table 3 we show the best-fitting parameter values from Full-Modeling fits with and without analytic marginalization. We see that the parameters that we marginalize over are well behaved and on the same order as they take when being sampled.

We also note the third term in eq. (A.5), $-(1/2)\log|L|$, which is the log of the determinant of the (linear parameter) part of the Fisher matrix. One prior choice that one can very easily implement is a "partial Jeffrey's prior" which removes this term from the likelihood. This prior can cause significant shifts in constraints in cases where parameter projection effects are noticeable, as the Jeffrey's prior removes some of the phase space volume from the likelihood. We discuss the implications of such a prior in appendix B.

B Parameter projection effects and the role of priors

In this section we discuss the role of priors on the parameters of our model and the effect they can have on parameter projection effects — defined here as shifts in the marginal posteriors away from the maximum likelihood regions due to a non-Gaussian posterior surface. These effects frequently arise when there are several parameters in the model that are poorly
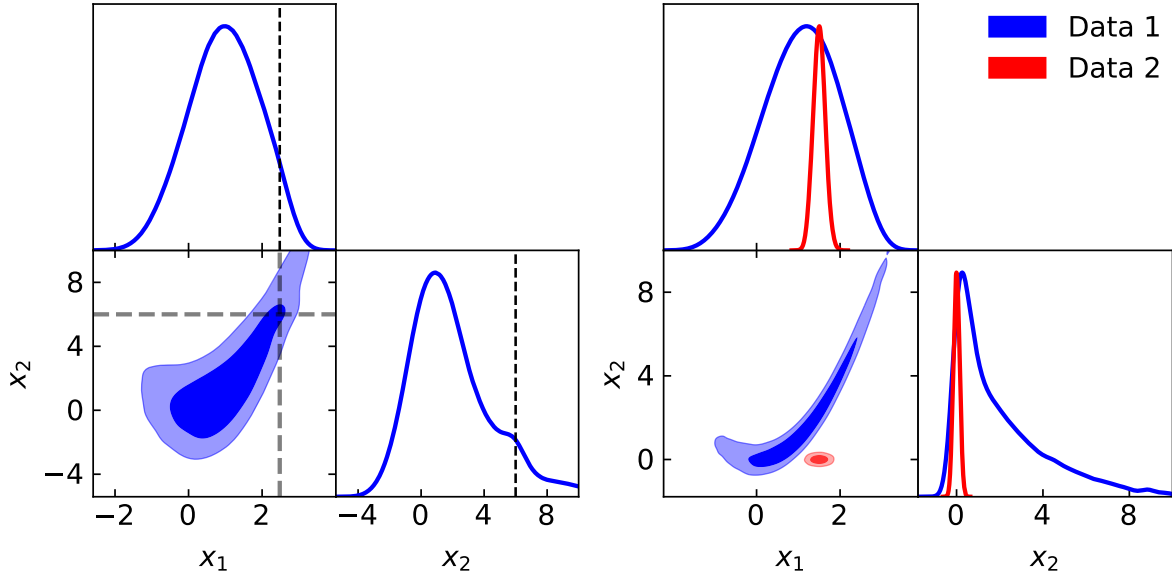
Figure 15. Toy model examples of information loss in projected posteriors. The left panel shows the posteriors from sampling from a likelihood distribution that is constructed out of the sum of a small Gaussian and a Rosenbrock function in 2D. The dashed lines label the maximum likelihood values of the two parameters x_1 and x_2. The 'truth', and likelihood maximum, appears to be in the tail of the 1D posteriors due to the large volume (area) at only slightly lower likelihood near $x_1 \approx x_2 \approx 0$. The right panel shows posteriors from two different "data sets" (different likelihood distributions). Data 1 is constructed from a Rosenbrock function and Data 2 is a Gaussian distribution. While in the full space (2D) it is clear the posteriors disagree, in projection (here 1D) the posteriors appear consistent.

constrained or partially degenerate. If there are degeneracies between parameters in the model, regions of the parameter space far from the maximum likelihood point may have very little likelihood penalty compared to the best fit. In spaces with large numbers of dimensions the "parameter volume" in such regions can be large, and integration over a subset of these parameters can shift the peaks or means of the marginal posterior distribution significantly away from the maximum likelihood values or the "input cosmology" in our tests. In addition, when the data are not sufficiently powerful the constraints on the cosmological parameters can depend on the choice of priors and the parameterization.

It is notoriously difficult to visualize complex probability distributions in high-dimensional spaces, and unfortunately projections *necessarily* remove information even if they are given from many viewpoints. For this reason marginal likelihoods can appear consistent (i.e. overlap in projection) when they are not and they can appear inconsistent when they are actually consistent. Even linear changes of the projection axes can change the appearance of concordance. Such issues are by no means specific to our models: projection effects in high-dimensional parameter spaces have been encountered in many areas of cosmology and have been widely discussed in the literature (see e.g. refs. [83–87] for recent discussions).

In figure 15 we show two toy model examples of projections, where the left plot is inspired by figure 1 of ref. [85] and the right plot is inspired by figure 1 of ref. [84]. For the first example, we construct a fake likelihood distribution by adding a Rosenbrock function, $f(x_1, x_2) = (1.0 - x_1)^2 + 0.5(x_2 - x_1^2)^2$, and a sharp 2D Gaussian centered at $(\bar{x}_1 = 2.5, \bar{x}_2 = 6)$

with a width of $\sigma = 0.25$ along both parameter directions. The maximum of the total likelihood distribution is very close to the center of the Gaussian, and is labeled with grey dashed lines in the figure. However, the contribution of the Rosenbrock function peaks at $(1.0, 1.0)$ but in a much more gradual way. The result is more likelihood "volume" for the MCMC to explore near $(1.0, 1.0)$ than near the true maximum of the whole likelihood. As a result, the marginal posterior distributions for parameters x_1 and x_2 are significantly offset from the true best-fitting points.

The second cautionary example of projections is presented in the right panel of figure 15 and shows posteriors from two "data sets", which we simulate by constructing two different fake likelihood distributions. For Data 1 we again use a Rosenbrock function, $f(x_1, x_2) = (1.0 - x_1)^2 + 10(x_2 - x_1^2)^2$ and for Data 2 we use a Gaussian with means $(\bar{x}_1 = 1.5, \bar{x}_2 = 0.0)$ and widths of 0.2. In this example we demonstrate how the constraints on x_1 and x_2 appear to agree for the two data sets when looking at the 1D posteriors, but in the 2D panel the two data sets are clearly in tension. This serves as a cautionary tale about interpreting constraints from a multi-dimensional posterior surface when looking at the projections onto lower dimensions. It is naturally difficult to visualize an N-dimensional volume, but looking only at 1D or 2D projections of the full distributions might lead one to misinterpret results.

Finally, as an honorable mention, we refer readers to figure 7 of ref. [83] in which the authors show a toy model of posteriors from two different data sets with three sampled parameters, x, y, z. The posteriors for these three parameters are consistent between data sets. However, after performing a linear transformation to new coordinates, $(x + y - z, x + z - y, y + z - x)$ one finds discrepant constraints on $x + y - z$. This shows that tensions can be hidden due to particular choices of parameterization, and that appropriate coordinate-independent metrics are necessary to measure the consistency between data sets or results.

B.1 Projection effects for DESI

To demonstrate the impact of projection effects in the specific case of DESI data with covariances similar to those expected from the first year we turn to synthetic data created with `velocileptors` for each of the seven DESI Y1 redshift bins: BGS $(0.1 < z < 0.4)$, LRG $(0.4 < z < 0.6, 0.6 < z < 0.8, 0.8 < z < 1.1)$, ELG $(0.8 < z < 1.1, 1.1 < z < 1.6)$, and QSO $(0.8 < z < 2.1)$. Since the data we are fitting to have been generated from the model, with no noise added, the best-fit point occurs at "truth" and has $\chi^2 = 0$. However χ^2 may rise slowly along some directions which have significant volume, shifting the marginalized posteriors away from the best-fit point. While the ΛCDM (with and without fixing n_s) and kΛCDM models do not exhibit significant projection effects, we do observe them for wCDM. We show the wCDM joint fits to the seven Y1 redshift bins in figure 16. Note that the marginal posteriors on several parameters (black lines in the left hand panels of figure 16) peak way from the input model, even though the model is, by construction, a good fit to the (mock) data and the maximum likelihood point is (again by construction) at the true values of the parameters. As the data become more constraining these projection effects are reduced — shown as the red contours in the same figure where the errors have been scaled down by a factor of 5. Note that some projection effects are still visible in the red contours. The posterior for Ω_m is still offset by a non-trivial fraction of its "new" error bar, but the absolute value of the
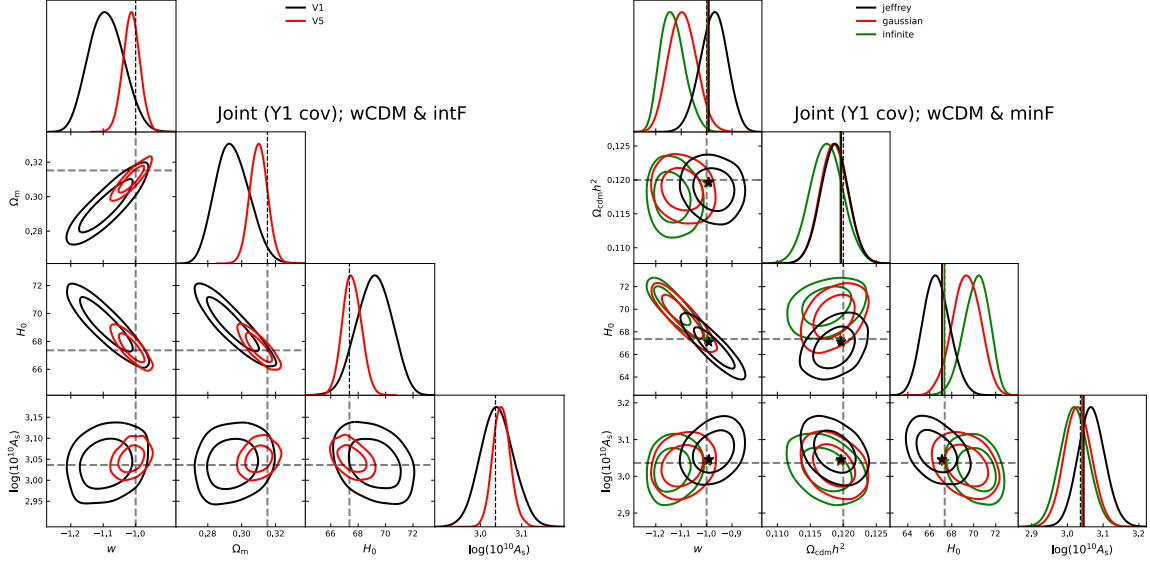
Figure 16. Full-Modeling wCDM fits to synthetic data created in each of the DESI Y1 redshift bins with corresponding analytic covariances. `Left:` intermediate fits with the real Y1 volumes (black) compared to fits with the covariances rescaled by a factor of 1/5. `Right:` minimal freedom fits with three different prior choices on the analytically marginalized counter and stochastic terms. The black contours correspond to a *partial* Jeffrey's prior (only on linear parameters) discussed in the text, the red contours show the fit with our usual Gaussian priors described in table 1, and the green contours correspond to 'infinite' priors. The stars in the 2D panels and solid vertical lines in the diagonal (1D posterior) panels denote the best-fit models obtained by running a minimizer starting at the MAP values of the chains.

offset is reduced. As we continue to reduce the error bars the contours shrink to eventually be δ-functions at the true values. It is also worth noting another feature of these projection effects. They typically occur when there are many parameters, some of which are partially degenerate. They also tend to lead to shifts that are $\mathcal{O}(1\,\sigma)$. This is because the likelihood falls as $\exp[-\chi^2/2]$ moving away from the best-fit point, while the volume in parameter space grows as a power of the "parameter distance". Eventually the Gaussian overcomes the impact of the volume. In the right panel of figure 16 we show wCDM constraints to the same synthetic data using three choices of priors on the linear parameters $(\alpha_0, \alpha_2, \mathrm{SN}_0, \mathrm{SN}_2)$: infinite uniform, Gaussian, and the (partial) Jeffrey's prior. The stars and solid vertical lines denote the best-fit values obtained from running a minimizer, and demonstrate that the shift between marginal posteriors and maximum likelihood values are due to projection effects. We find that these projection effects are slightly reduced when switching from the flat to Gaussian prior, showing that the Gaussian priors on the linear parameters are not entirely uninformative. The projection effects are more significantly reduced when applying the Jeffrey's prior and we discuss the implications of using such a prior in the next section.

B.2 Jeffrey's prior and reparameterizations

In addition to shifts in the posteriors such that they peak away from the 'true' values, insufficiently constraining data in a high-dimensional parameter space can lead to increased

sensitivity to priors and choice of parameterizations. This is another manifestation of the likelihood not dominating the posterior and is a generic feature of inference in high dimensions. If we had firm theoretical reasons to prefer one model parameterization over another this would not be a problem, but in practice there are several choices between which there is little theoretical preference. We discuss some of these implications here — first discussing the choice of parameters and then the Jeffrey's prior.

A natural[15] set of parameters for the model would be the cosmological parameters (e.g. σ_8) and the bias parameters and counterterms (b_i and α_i). However some of these are at least partially degenerate. Lowering A_s or σ_8 while raising α can leave $\alpha\,k^2 P$ unchanged, and a similar upward adjustment of b_i can reduce much of the impact from the other terms so that χ^2 changes little. Since, for linear priors on b_i and α_i, there is more "volume" at large values than small there is a natural tendency to shift the posterior to lower σ_8. The quantities best-constrained from observation are the power spectrum multipoles, and in particular the monopole. For this reason we use parameters that are closer to the data space, i.e. $b\sigma_8$ rather than b (see table 1). While this is a natural choice, in terms of the b_i it corresponds to a prior that rises with σ_8 [88]. For example, the Jacobian translating between (b, σ_8) and $(b\sigma_8, \sigma_8)$ is simply σ_8. Inference using the second set of parameters is thus equivalent to inference using the first, plus a prior $P(\sigma_8) \propto \sigma_8$. When σ_8 is not well constrained by the data, this prior choice will shift the marginal posterior. Similar comments hold for the other parameters of course.

A method that is sometimes used in the statistics literature to reduce the impact of parameter changes is to include a "Jeffrey's prior". This corresponds to the square root of the determinant of the Fisher matrix, and has the same role as the familiar $\sqrt{-g}\,d^4x$ in General Relativity. If implemented consistently, this removes the Jacobian from transformations of variables and so is sometimes termed[16] "uninformative". There are some concerns about taking this approach in our situation however.[17] First, we do not believe that the physics indicates that e.g. $(\ln[1 + b^{10}], \cosh \sigma_8)$ is as good a parameter set as $(b\sigma_8, \sigma_8)$ for example. Our parameters have at least some theoretical justification that we'd like to include as "prior information" in our model specification. Secondly, as usually implemented, the Jeffrey's prior is a strong function of several key cosmological parameters.

To see this, let us consider the *partial* Jeffrey's prior that is sometimes introduced. This involves computing $\sqrt{\det F}$ for only those parameters that enter the model linearly (if all parameters enter linearly, then this is the "full" Jeffrey's prior, however in that limit the likelihood is Gaussian so the issue of projection effects does not arise). The calculation in the previous appendix shows that introducing such a prior is equivalent to dropping the $\log ||L||$ term in eq. (A.5) (see also ref. [86]), making this a very easy change to make. That this prior is a strong function of the underlying cosmological parameters is most easily seen by

---

[15]This is not the only choice. One could imagine choosing e.g. log priors in the mass scale of the halos hosting the galaxies, or linear deviations from the peak-background split prediction (where the $b_{n>1}$ are non-linear functions of b_1), or many other choices.

[16]While common, this nomenclature is incorrect. A much better term would be "reparametersation invariant" since in general — and in our case — the prior is "informative" from the point of view of inference.

[17]The Jeffrey's prior and problems with it are also discussed in ref. [89], including an example from ref. [90].
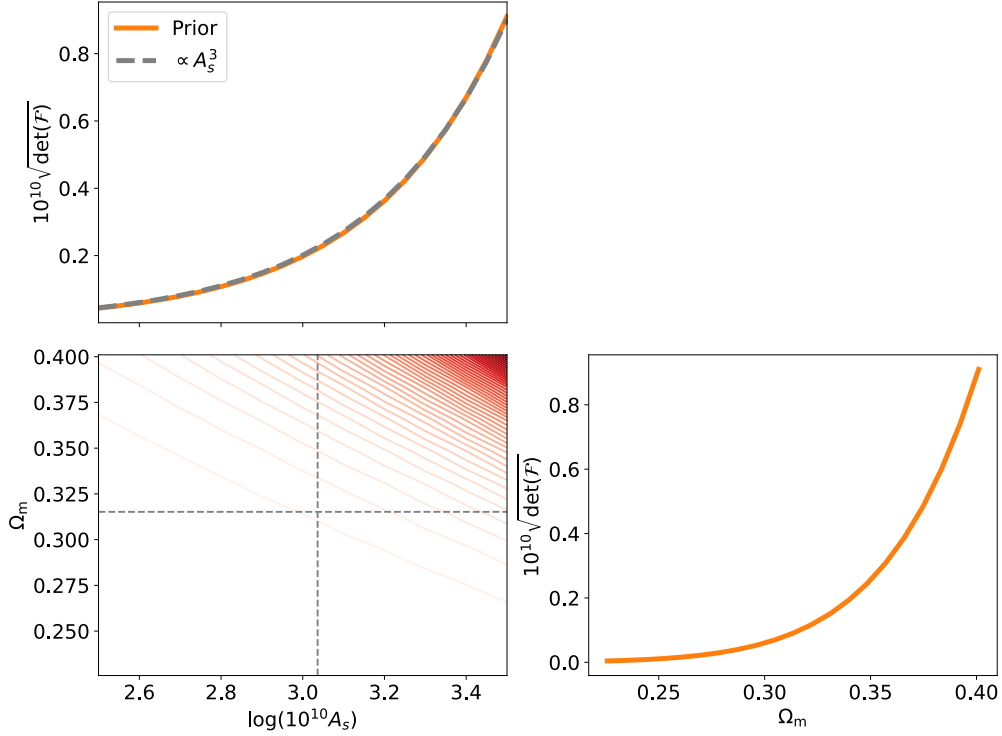
Figure 17. A 2D slice through the *partial* Jeffrey's prior for ΛCDM. We show the variation of the prior in the Ω_m and A_s directions, with all other parameters fixed to their best fit. The dashed grey lines in the lower left panel show the values of $\log(10^{10}A_s)$ and Ω_m along which $\det F$ was evaluated in each of the other panels. The dashed grey line in the upper left panel shows a power law $\propto A_s^3 \propto \sigma_8^6$. Note the strong dependence of this prior on the cosmological parameters (see text).

again considering σ_8. The Fisher matrix has the form

$$F \sim \frac{\partial(\text{theory})}{\partial(\text{param})} C^{-1} \frac{\partial(\text{theory})}{\partial(\text{param})} \sim (\text{template}) \, C^{-1} \, (\text{template}) \tag{B.1}$$

where in the second step we have used the fact that for parameters entering linearly the derivative is just some linear-parameter-independent template — e.g. for $\alpha k^2 P$ it would be $k^2 P$. In the case of our perturbative model, each of these 'templates' is P_{lin} or some integral over one or more powers of P_{lin} and thus we expect the template to scale as a power of A_s or σ_8. The Fisher matrix is thus also a (high) power of A_s or σ_8 and so including such a prior has the effect of shifting the marginal posterior to higher σ_8.

Figure 17 shows a 2D slice through this (high-dimensional) prior to illustrate the previous points. We have chosen to show the variation in the Ω_m and A_s directions with all of the other parameters held fixed at their best-fit points. The strong dependence on A_s is clear ($\propto A_s^3 \propto \sigma_8^6$), and has been described above. The Ω_m dependence can be understood similarly. Raising Ω_m, with all other parameters fixed, changes the shape of P_{lin} with more power on the quasi-linear scales of relevance to DESI (and less power at large scales). The increase in the amplitude of P_{lin} increases $\det F$ in the same manner as for A_s or σ_8. The dependence on each of the other parameters can be similarly computed and understood, though they are not shown here for simplicity. The introduction of such a prior is thus "informative" or "strongly

informative" in the sense of introducing non-negligible shifts in the marginal posteriors given the size of the uncertainties. We note that in making figure 17 we used the more traditional form for the counterterms, e.g. $\alpha k^2 P_{\text{lin}}$ instead of the parameterization of eq. (3.6), since it is in that context that (partial) Jeffrey's priors have typically been discussed. For most of this paper we have chosen parameters scaling like $\alpha\sigma_8^2$, meaning that the "template" is closer to $k^2 P_{\text{lin}}/\sigma_8^2$ and is therefore largely independent of σ_8. Indeed, we find that in our preferred parameterization the (partial) Jeffrey's prior scales much more weakly with σ_8 than what is usually encountered. However, the strong dependence on Ω_{m} and other cosmological parameters is unaffected by this particular reparameterization.

There are two things to note about these examples. First, in each case the shift in the marginal posterior was accomplished by the introduction of a what is effectively a prior, and not by any change in the model or the data. It relies on the fact that the data are not sufficiently constraining such that such prior or parameterization choices are relevant. Second, the two approaches change the prior through different parts of the theory. In the first case we modified the biases while in the second we introduced a prior through the counterterms.

Luckily the existing theoretical models are sufficiently accurate to model much more constraining data than DESI Y1 without the need to introduce additional free parameters (see the main body of the paper and refs. [38, 47]). As the data become more constraining the impact of parameter choices and priors is expected to reduce, as shown earlier. Combining the DESI data with other datasets that can break degeneracies is also expected to reduce the impact of these effects. In this sense, the Y1 data may well be a "worst case" scenario.

C Connection to the halo model

It is sometimes helpful to establish the expected sizes of the terms in the theoretical model. This can be done through arguments of self-consistency (see main text), and by comparing to other models. In this appendix we compare the PT approach to a simplified, analytical halo model [91, 92] with the goal of understanding the expected size of the stochastic terms (see also the discussion in ref. [48]). Since our goal is to gain insight, we shall deal with an analytically tractable version of the halo model in which galaxies reside in spherical, self-similar halos whose centers are distributed according to biased linear theory with scale-independent bias. If $n(M)$ is the volume density of halos per unit mass, and each halo has a Fourier-space density profile $u(k, m, z)$, normalized to unity as $k \to 0$, then the power spectrum is (see e.g. ref. [93] for a recent, pedagogical discussion with references to the original literature)

$$P_g(k, \mu, z) = P_g^{2-\text{halo}}(k, \mu, z) + P_g^{1-\text{halo}}(k, \mu, z) + P_g^{\text{shot}}. \tag{C.1}$$

If $N_{\text{cen}}(m)$ and $N_{\text{sat}}(m)$ denote the mean number of centrals and satellites in a halo of mass m the mean number density of galaxies is simply $\bar{n}_g = \int dm\, n(m)\left[N_{\text{cen}}(m) + N_{\text{sat}}(m)\right]$. To compute the clustering we need to know the statistics of the galaxy occupation, and we shall follow standard practice in assuming the centrals are Bernoulli distributed while the satellites are Poisson distributed.

Under the above assumptions the 2-halo term in the power spectrum is given by:

$$P_g^{2\text{-halo}} = \left(b_g + F\mu^2\right)^2 P_{\text{lin}}, \tag{C.2}$$

where the bias is

$$b_g(k, \mu, z) \equiv \frac{1}{\bar{n}_g} \int dm \, n(m) \, b(m) \left[N_{\text{cen}} + N_{\text{sat}} u(k, m, z) e^{-k^2 \mu^2 \sigma_d^2(m)/2} \right] \quad \text{(C.3)}$$

and the effective growth rate of structure is

$$F(k, \mu, z) \equiv f \int dm \, n(m) \left(\frac{m}{\bar{\rho}} \right) u(k, m) e^{-k^2 \mu^2 \sigma_d^2(m)/2} \quad \text{(C.4)}$$

which tends to f as $k \to 0$. In the above we have written the (linear) bias of a halo of mass m as $b(m)$ and the mean matter density in the Universe as $\bar{\rho}$. We have also used the fact that in going into redshift-space, the density profile acquires a damping factor from the virial motions in halos:

$$u(k, m, z) \to u(k, m, z) e^{-k^2 \mu^2 \sigma_d^2/2}, \quad \text{(C.5)}$$

where $\sigma_d^2(m)$ is the velocity dispersion of such a halo in distance units. The 1-halo term has in its integrand the term $\langle N(N-1) \rangle$ which, when expanded is:

$$\begin{aligned}
\langle N(N-1) \rangle &= \langle (N_{\text{cen}} + N_{\text{sat}})(N_{\text{cen}} + N_{\text{sat}} - 1) \rangle \\
&= \langle N_{\text{cen}}^2 - N_{\text{cen}} + 2N_{\text{cen}}N_{\text{sat}} + N_{\text{sat}}(N_{\text{sat}} - 1) \rangle \\
&= 2\langle N_{\text{cen}}N_{\text{sat}} \rangle + \langle N_{\text{sat}}(N_{\text{sat}} - 1) \rangle \\
&= 2\langle N_{\text{cen}} \rangle \langle N_{\text{sat}} \rangle + \langle N_{\text{sat}} \rangle^2 \quad \text{(C.6)}
\end{aligned}$$

where in going from the second to third line we used that $N_{\text{cen}} = 0, 1 \to \langle N_{\text{cen}}^2 \rangle = \langle N_{\text{cen}} \rangle$. We obtain the last equality by assuming that the centrals and satellites are uncorrelated and that N_{sat} follows a Poisson distribution, such that $\langle N_{\text{sat}}^2 \rangle = \langle N_{\text{sat}} \rangle^2 + \langle N_{\text{sat}} \rangle$. Using this, the 1-halo term becomes (dropping the $\langle \rangle$'s for simplicity):

$$P_g^{\text{1-halo}} = \frac{1}{\bar{n}_g^2} \int dm \, n(m) \left[N_{\text{sat}}^2 |u(k, m, z)|^2 e^{-k^2 \mu^2 \sigma_d^2(m)} + 2N_{\text{sat}}N_{\text{cen}} u(k, m, z) e^{-k^2 \mu^2 \sigma_d^2(m)/2} \right]. \quad \text{(C.7)}$$

Finally, the shot noise power spectrum is simply $P_g^{\text{shot}} = \bar{n}_g^{-1}$ if we assume Poisson fluctuations for the galaxies and halos.

Our perturbative model should be able to describe any 'complete' model of galaxy clustering, whether or not that model is correct in detail. We can make the connection by considering the low-k limit of the halo model. To make our expressions slightly simpler we shall make an additional approximation that $u(k, m, z) \approx 1$ on the scales of interest, which corresponds to assuming that $k r_{\text{vir}} \ll 1$. We shall further assume that $\sigma_d > r_{\text{vir}}$ so that the impact of virial velocities is more important than the fact that the satellites do not sit at the halo center. Under these approximations, and for small k,

$$b_g(k, \mu, z) \simeq \frac{1}{\bar{n}_g} \int dm \, n(m) \, b(m) \left[N_{\text{cen}} + N_{\text{sat}} \left(1 - \frac{1}{2} k^2 \mu^2 \sigma_d^2(m) \right) \right] \quad \text{(C.8)}$$

$$= \frac{1}{\bar{n}_g} \int dm \, n(m) \, b(m) N_{\text{gal}} - \frac{1}{2} k^2 \mu^2 \frac{1}{\bar{n}_g} \int dm \, n(m) \, b(m) N_{\text{sat}} \sigma_d^2(m) \quad \text{(C.9)}$$

$$= b_{\text{eff}} \left(1 - \frac{1}{2} k^2 \mu^2 \sigma_{2,\text{eff}}^2 \right) \quad \text{(C.10)}$$

The $k^2\mu^2$ term above, combined with the b or $f\mu^2$ term from the other power of b_g in eq. (C.2) contributes to the counterterms, α_i.

Since the mass-integral in F extends all the way to $m = 0$, the $k^2\mu^2$ correction is smaller than for the bias and we shall neglect it, taking $F \to f$ henceforth. The 1-halo term becomes

$$P_g^{\text{1-halo}} \simeq \frac{1}{\bar{n}_g^2} \int dm\, n(m)\, N_{\text{sat}} e^{-k^2\mu^2\sigma_d^2/2} \left[N_{\text{cen}} + N_{\text{sat}} e^{-k^2\mu^2\sigma_d^2/2} \right] \tag{C.11}$$

$$\approx \frac{f_{\text{sat}}}{\bar{n}_g} \left(\cdots - \frac{1}{2}k^2\mu^2\sigma_{1,\text{eff}}^2 + \cdots \right) \tag{C.12}$$

Thus we see that the halo model predicts that the stochastic terms are of order $\text{SN}_0 \sim \bar{n}_g^{-1}$ (from P_g^{shot} in eq. (C.1)) and $\text{SN}_2 \sim f_{\text{sat}}\sigma_{1,\text{eff}}^2/\bar{n}_g$ (from eq. (C.12)) as described in the main text. Here f_{sat} is the satellite fraction such that $f_{\text{sat}}\sigma_{1,\text{eff}}^2$ is the mean velocity dispersion of halos weighted by $N_{\text{cen}}N_{\text{sat}}$, such that roughly speaking $\sigma_{1,\text{eff}}^2$ is the mean velocity dispersion of the satellites in question. We often refer to $f_{\text{sat}}^{1/2}\sigma_{1,\text{eff}}$ as a "characteristic halo velocity" for simplicity.

The simple derivation above neglects several physical effects, including halo compensation and exclusion, correlations between the halo density and velocity profiles and between local environment and profile, correlations between mass bins in the halo shot noise, etc. It is sufficient for order of magnitude estimates, since most of the neglected effects also have characteristic size set by the mean inter-galaxy separation or the virial or infall velocity of the halo but it should not be taken as a 'complete' model of clustering. As a single example of an effect missed by this simple treatment, let us further consider the effect of virial motions in eq. (C.5). Another way to account for the effect of FoG in the galaxy power spectrum is to introduce a random velocity field $\epsilon_i(s)$ to each galaxy, such that the observed position is $s + \hat{n} \cdot \epsilon_{\hat{n}}$. In this case the galaxy 2-point function with these additional velocities is [94, 95]

$$P(k,\mu) = \int d^3s\, e^{i\boldsymbol{k}\cdot\boldsymbol{s}} \left\langle e^{ik\mu(\epsilon_{\hat{n}}(\boldsymbol{s})-\epsilon_{\hat{n}}(\boldsymbol{0}))} \big(1 + \delta_g(\boldsymbol{s})\big)\big(1 + \delta_g(\boldsymbol{0})\big) \right\rangle$$

$$\approx \int d^3s\, e^{i\boldsymbol{k}\cdot\boldsymbol{s}} \left\langle e^{ik\mu(\epsilon_{\hat{n}}(\boldsymbol{s})-\epsilon_{\hat{n}}(\boldsymbol{0}))} \right\rangle \big(1 + \xi_g(\boldsymbol{s})\big) \tag{C.13}$$

where in the second line we have made the (unphysical) assumption that the virial motions and galaxy densities are uncorrelated in order to isolate the pure effect of virial velocities usually called FoGs (in the literature models making this approximation are frequently referred to as "dispersion" or "streaming" models). The expectation value of the exponential can be expanded in powers of $k\mu$ as

$$\ln \left\langle e^{ik\mu(\epsilon_{\hat{n}}(\boldsymbol{s})-\epsilon_{\hat{n}}(\boldsymbol{0}))} \right\rangle = 1 - k^2\mu^2 \left[\sigma_v^2 - \xi_\epsilon(\boldsymbol{s}) \right] + \mathcal{O}(k^3\mu^3), \tag{C.14}$$

where ξ_ϵ is the correlation function of the virial velocities projected along the line of sight. Since it describes virial motions, this correlation must fall rapidly to zero outside of the halo radius, R_h, and asymptote to the mean square velocity, σ_v^2, as $s \to 0$. Expanding this cumulant to first order we see that, in addition to the damping of the profile coming from $-k^2\mu^2\sigma_v^2$ in eq. (C.14) we also gain the contribution

$$P(k,\mu) \supset k^2\mu^2 \int d^3s\, e^{i\boldsymbol{k}\cdot\boldsymbol{s}}\, \xi_\epsilon(\boldsymbol{s})\big(1 + \xi_g(\boldsymbol{s})\big) \approx k^2\mu^2(1 + \sigma_g^2) \int d^3s\, e^{i\boldsymbol{k}\cdot\boldsymbol{s}}\, \xi_\epsilon(\boldsymbol{s}) \tag{C.15}$$

where we have used that the linear galaxy density is smooth compared to the support of ξ_ϵ and σ_g^2 is its the mean on the halo scale. The integral in the final expression is simply the noise spectrum of the virial motions, which we expect to be positive and white on large $(> R_h)$ scales and of order $\sim \sigma_v^2 R_h^3$. In order to differentiate between satellites and centrals we can simply set $\epsilon = 0$ for central galaxies such that the cumulant in equation (C.14) is instead simply unity for the central-central correlation and $1 - \frac{1}{2}k^2\mu^2\sigma_v^2$ for the central-satellite cross correlation. This gives the FoG prescription in the 'analytic halo model', derived above, with the addition of a positive, scale-dependent noise along the line of sight.[18]

We reiterate that our aim here was to motivate the scale of stochastic contributions and not to make claims about what numerical value (or even sign) they will take. We see that the term discussed above, while missed by the halo model, did scale in the same manner as the included terms as we stated above. Other allowed parameter combinations, such as $R_h^4\sigma_v$ for the stochastic piece, should be subdominant.

D Further tests

D.1 Dependence on $\omega_{\rm b}$ prior

We next test the dependence of our constraints on the prior set on $\omega_{\rm b}$. The standard setting that we choose is a Gaussian prior centered on $\omega_{\rm b}^{\rm true} = 0.02237$ with a width of $\sigma = 0.00037$, which is based on the recentmost Big-Bang Nucleosynthesis (BBN) constraints on primordial deuterium abundance [67] which places stringent constraints on $\omega_{\rm b}$. We test the dependence on the prior by loosening it to $\sigma = 0.001$. The results are shown in figure 18. Within each individual method we show results for the covariance appropriate to the single-box volume. We find that for all three methods, H_0 becomes significantly less constrained. Meanwhile the $\Omega_{\rm m}$ constraints remain unchanged in all methods.

In the Full-Modeling analysis, the measurement of $\Omega_{\rm m}$ is extracted from the shape of the power spectrum and scale of matter-radiation equality $k_{\rm eq}$, and these depend on the full matter abundance rather than ω_b and $\omega_{\rm cdm}$ separately. We thus do not see a degradation in the $\Omega_{\rm m}$ constraint when the prior on ω_b is relaxed. In the template and ShapeFit analyses $\Omega_{\rm m}$ is inferred from the compressed parameters, and because $f \simeq \Omega_{\rm m}^{0.55}$ we can extract a measurement of $\Omega_{\rm m}$ from the compressed amplitude parameter without any dependence on ω_b prior. In the ShapeFit case, additional constraining power on $\Omega_{\rm m}$ comes from the shape parameter m, but just like in the Full-Modeling case this power spectrum shape information translates to a measurement $\Omega_{\rm m}$ without any reliance on ω_b specifically.

For the H_0 measurement we do observe a significant degradation in constraining power when the prior on ω_b is relaxed. In the template analysis, information about cosmological distances is extracted from the BAO feature and thus constrains $H(z)r_d$ and $D_A(z)/r_d$. Breaking the degeneracy between H_0 and r_d requires a physical (dimensionful) length scale for the distance-redshift relation beyond just the angular size of the BAO feature [96]. This is accomplished with knowledge about ω_b (which determines r_d) from either BBN or CMB and then leads to a direct measurement of H_0. Therefore, relaxing the prior on ω_b worsens the constraint on H_0. The inclusion of the shape parameter m, while in general improving

---

[18]We thank Misha Ivanov for pointing out that the sign of this effect in N-body simulations is often positive.
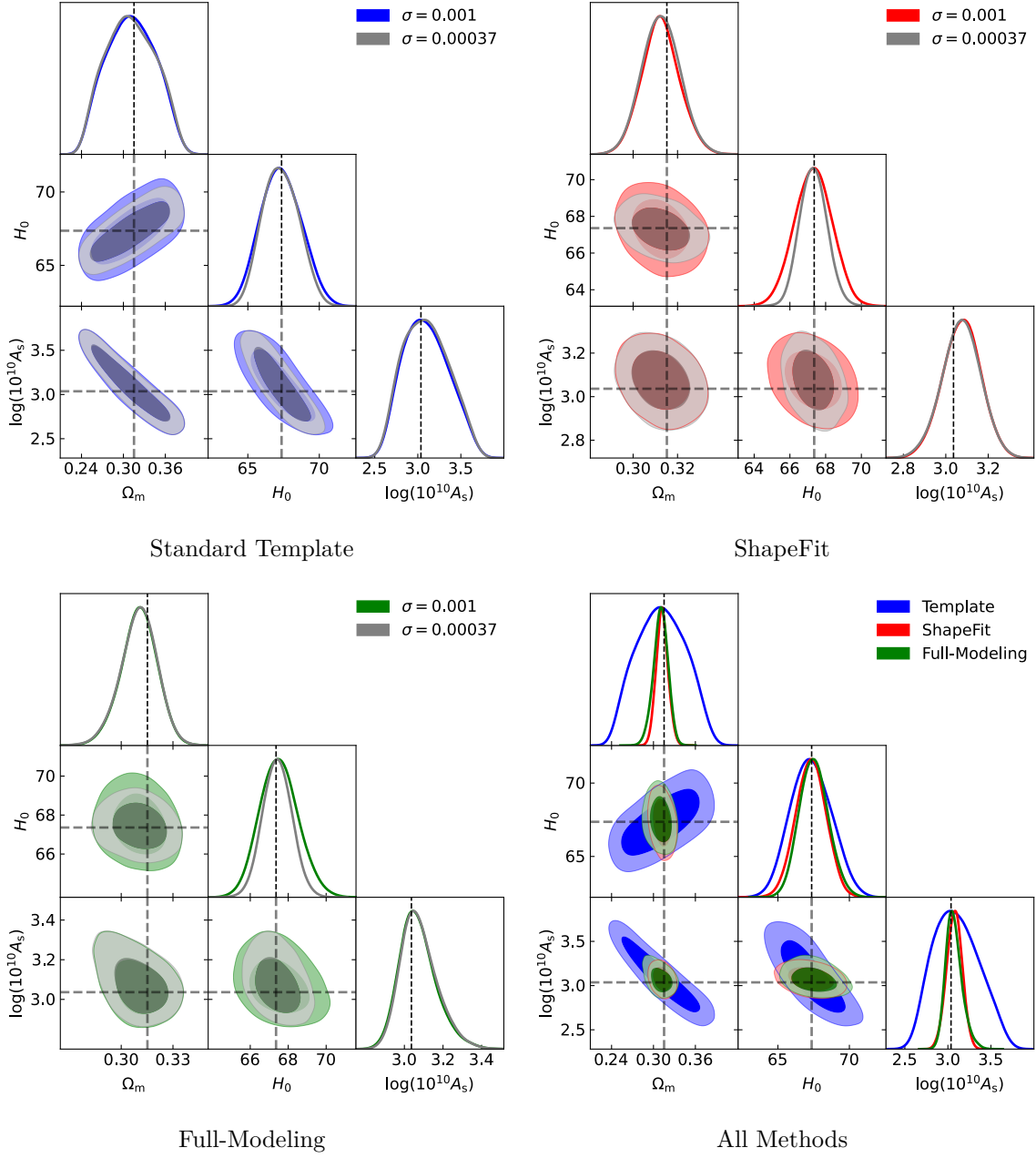
Figure 18. Comparison of constraints when loosening the prior on ω_b from $\sigma = 0.00037$ to $\sigma = 0.001$. In all cases we use the single box covariance. The bottom right plot shows a comparison of the three modeling methods while using a $\sigma = 0.001$ prior on ω_b.

constraints when compared to the standard template, does not compensate for the changes in ω_b information and therefore ShapeFit also experiences worse H_0 constraint. The Full-Modeling method can in principle constrain ω_b (and by extension r_d) in the absence of an external prior because the amplitude of BAO wiggles depend on ω_b and $\omega_{\rm cdm}$ and can be modulated in Full-Modeling analyses, but this is still a much weaker constraint than what can be accomplished with a BBN prior [97].

D.2 Minimal and maximal freedom in the bias parameters

In this section we discuss three possible choices in freedom in the bias parameters. In total there are four bias parameters: b_1, b_2, b_s, and b_3. The first two parameters multiply the initial $\delta_0(q)$ and $\delta_0^2(q)$ overdensity fields in the bias expansion. The non-local tidal bias parameter, b_s multiplies the initial shear field and, due to degeneracies between terms, the third order bias contributions are combined into a single operator with coefficient b_3. In the Lagrangian picture the bias contributions are evaluated at the initial positions $\boldsymbol{q}$, whereas in the Eulerian framework the bias expansion is performed at observed coordinates $\boldsymbol{x}$. This implies that the non-local bias terms in Eulerian PT are dependent on both the initial Lagrangian non-local contributions as well as gravitational evolution such that the Eulerian biases are affine transformations of the Lagrangian ones, with coefficents dependent on the definition of the bias operators in each space. Therefore, one commonly sees in the literature of Eulerian PT models (e.g. [51, 98]) a "minimal" and "maximal" freedom parametrization where the first assumes a local Lagrangian bias initially with no third-order contributions ($b_s^L = b_3^L = 0$) and that tidal and 3rd order biases are induced entirely by gravitational nonlinearity [99]. In such a case, the tidal and third order Eulerian biases would coevolve with the linear bias terms, i.e. $b_i^E \propto b_1^L = b_1^E - 1$. In the maximal freedom case, on the other hand, all bias parameters are allowed to vary independently.

The two other Fourier space EFT models that will be used in the DESI collaboration, FOLPSν and PyBird, are both based on the Eulerian frameworks and it has been shown that `velocileptors`LPT and EPT agree closely with the other two models under a consistent choice of parametrization [47]. For this reason we are interested in comparing the three parameter choices within LPT. In the Lagrangian picture, it is not clear how well motivated the initially local bias assumption is, and for most of this paper we chose an intermediate option in which the tidal bias b_s is allowed to vary along with b_1 and b_2, but the third order bias is kept fixed to zero, both because the cubic bias is expected to be small for intermediate mass halos and, more importantly, quite degenerate with the counterterms. We advise caution against restricting the parameter space further when fitting the high volume simulations with the 25 box covariance, as the tightness of the error bars can result in poor behavior of the model, which we demonstrate in the left panel of figure 19. While at $k_{\rm max} = 0.18\,h{\rm Mpc}^{-1}$ the constraints are fine, raising the scale cut to $k_{\rm max} = 0.2\,h{\rm Mpc}^{-1}$ results in a bimodal distribution appearing in the posteriors, most likely driven by some two-loop effects. However, including the b_s parameter fixes the bimodal behavior and we instead recover more Gaussian posteriors. We also show that this problem is induced by the extremely tight covariance from the full 25- cubic box volume. In the right panel of figure 19 we compare the Full-Modeling constraints between both $k_{\rm max}$ values with minimal freedom for the single box volume and find the two in agreement without any non-Gaussian behavior.

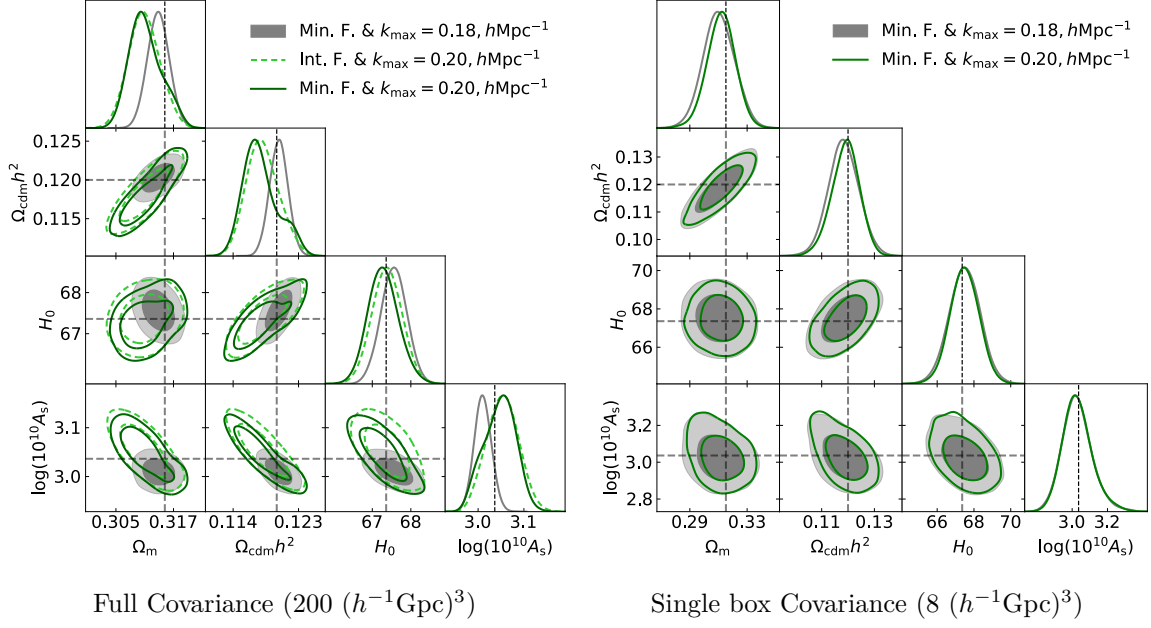Full Covariance (200 $(h^{-1}\text{Gpc})^3$) Single box Covariance (8 $(h^{-1}\text{Gpc})^3$)

Figure 19. *(left):* Full-Modeling constraints for the minimal freedom parametrization with $k_{\max} = 0.18\,h\text{Mpc}^{-1}$ (grey), $k_{\max} = 0.20\,h\text{Mpc}^{-1}$ (dark green), and intermediate freedom case with $k_{\max} = 0.20\,h\text{Mpc}^{-1}$ (light green dashed). In the minimum freedom case there is a bimodal distribution (most pronounced in $\Omega_{\text{cdm}}h^2$) that appears when $k_{\max}$ is raised from 0.18 to 0.2 $h\text{Mpc}^{-1}$. The bi-modality disappears if the b_s parameter is included, as in the intermediate freedom case. *(right):* Full-Modeling constraints in the minimal freedom case with $k_{\max} = 0.18$ (grey) and $0.20\,h\text{Mpc}^{-1}$ (green) using the single-box covariance.

Choosing the single-box covariance and a $k_{\max} = 0.2\,h\text{Mpc}^{-1}$ we proceed with the comparison between the minimal, intermediate, and maximal freedom bias parametrizations. The results are shown in figure 20 for the Full-Modelling and ShapeFit methods. We find that the parameters primarily controlling the shape of the linear power spectrum, i.e. Ω_{m} in FM and m in SF, are the most affected by the differences in parameterization. Meanwhile the amplitude σ_8 in FM is fairly resistant to these changes. We remind the reader that σ_8 is more directly constrained in LSS analyses than $\log(10^{10}A_{\text{s}})$, suggesting it is a better way of quoting the normalization of the theory for these purposes. We find that fixing $b_3 = 0$ does not result in significant offsets away from the true cosmology, and mostly just tighten constraints. This is consistent with previous tests on the bias parametrization, and our standard choice of fixing b_3 in this paper mirrors that of previous analyses using `velocileptors` [48, 88]. We conclude this section by reiterating that despite the improvement in constraining power obtained in the minimal freedom case, fixing both b_s and b_3 can lead to poor performance of the model in capturing the nonlinear effects that become increasingly important at very high simulation volumes, and it therefore is safer to use the intermediate freedom choice. In addition, depending on the method of galaxy sample-selection, larger values of b_s than expected can occur due to assembly bias (see e.g. ref. [100]). This further motivates keeping b_s as a free parameter. While we have justification for the choice of fixing $b_3 = 0$, it is also a valid and more conservative option to allow b_3 to vary and we do not strongly discourage the maximal freedom choice in future analyses.
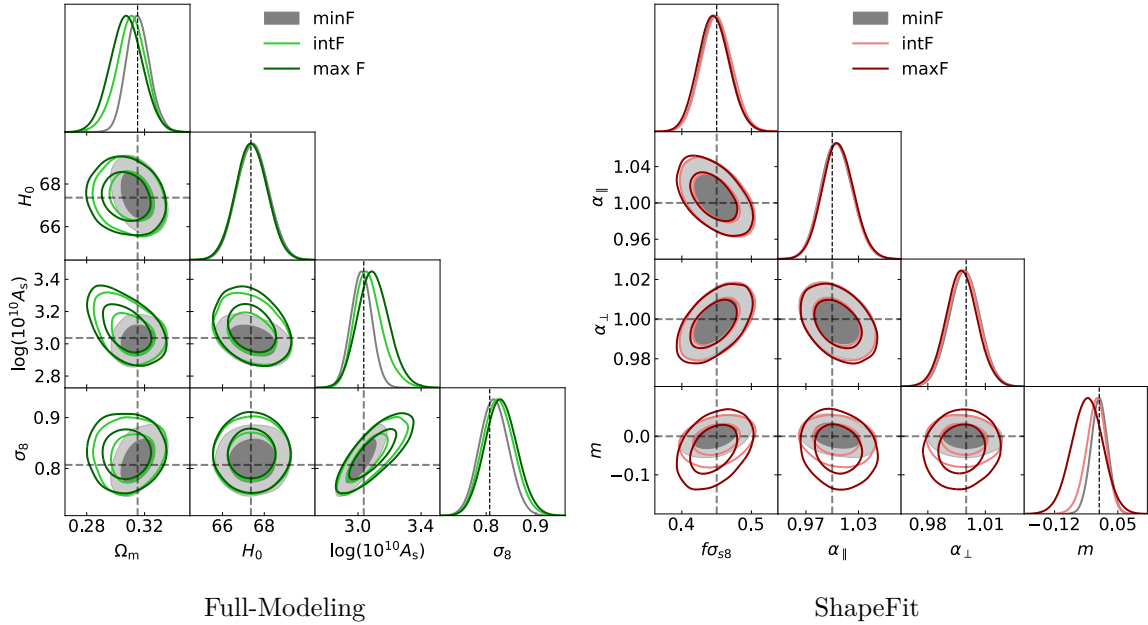
Figure 20. Comparison of minimal ($b_s = b_3 = 0$), intermediate ($b_3 = 0$), and maximal (b_s, b_3 free) freedom parametrizations using the single box covariance and $k_{\max} = 0.20\,h\mathrm{Mpc}^{-1}$. The Full-Modeling constraints are in the left figure and ShapeFit on the right.

D.3 Including hexadecapole

The 1-loop LPT model we use predicts the full angular dependence of the power spectrum $P(k, \mu)$ and therefore makes consistent predictions for the power spectrum hexadecapole and above in addition to the monopole and quadrupole. However, it should be noted that since the linear theory hexadecapole is substantially smaller than the monopole or quadrupole (there are no linear theory $\ell > 4$ multipoles) these higher multipoles will be more sensitive to nonlinear effects (e.g. Finger of God (FoG)), and thus the range of scales over which their 1-loop PT predictions is valid may be smaller. We present results of including the hexadecapole in figure 21 for the covariance of the single-box volume. We find a slight tightening of the constraints when including the hexadecapole.

In figure 22 we show in the left panel the ΛCDM parameter constraints of all three methods when fitting $\ell = 0, 2, 4$ instead of just $\ell = 0, 2$, using the covariance for the 25 box volume. As with the previous comparisons between methods, we find consistent constraints between ShapeFit and Full-Modeling and looser constraints for the standard template. We also test the dependence of the hexadecapole on it's k-range by lowering the upper bound from $k_{\max} = 0.2\ h\mathrm{Mpc}^{-1}$ down to 0.15 and 0.1 $h\mathrm{Mpc}^{-1}$, while keeping the range of scales of the monopole and quadrupole moments fixed at $0.2\ h\mathrm{Mpc}^{-1}$. While we see very little change in constraints in this case, other data sets may have significantly larger FoG effects (or observational systematics) that could affect the hexadecapole at $k \gtrsim 0.1\ h\mathrm{Mpc}^{-1}$. For this reason we still suggest using $k_{\max} = 0.1\ h\mathrm{Mpc}^{-1}$ for the hexadecapole and correspondingly widening the α_4 prior to $\mathcal{N}[0, 20]$ to maintain the 20% scaling at the new $k_{\max}$.
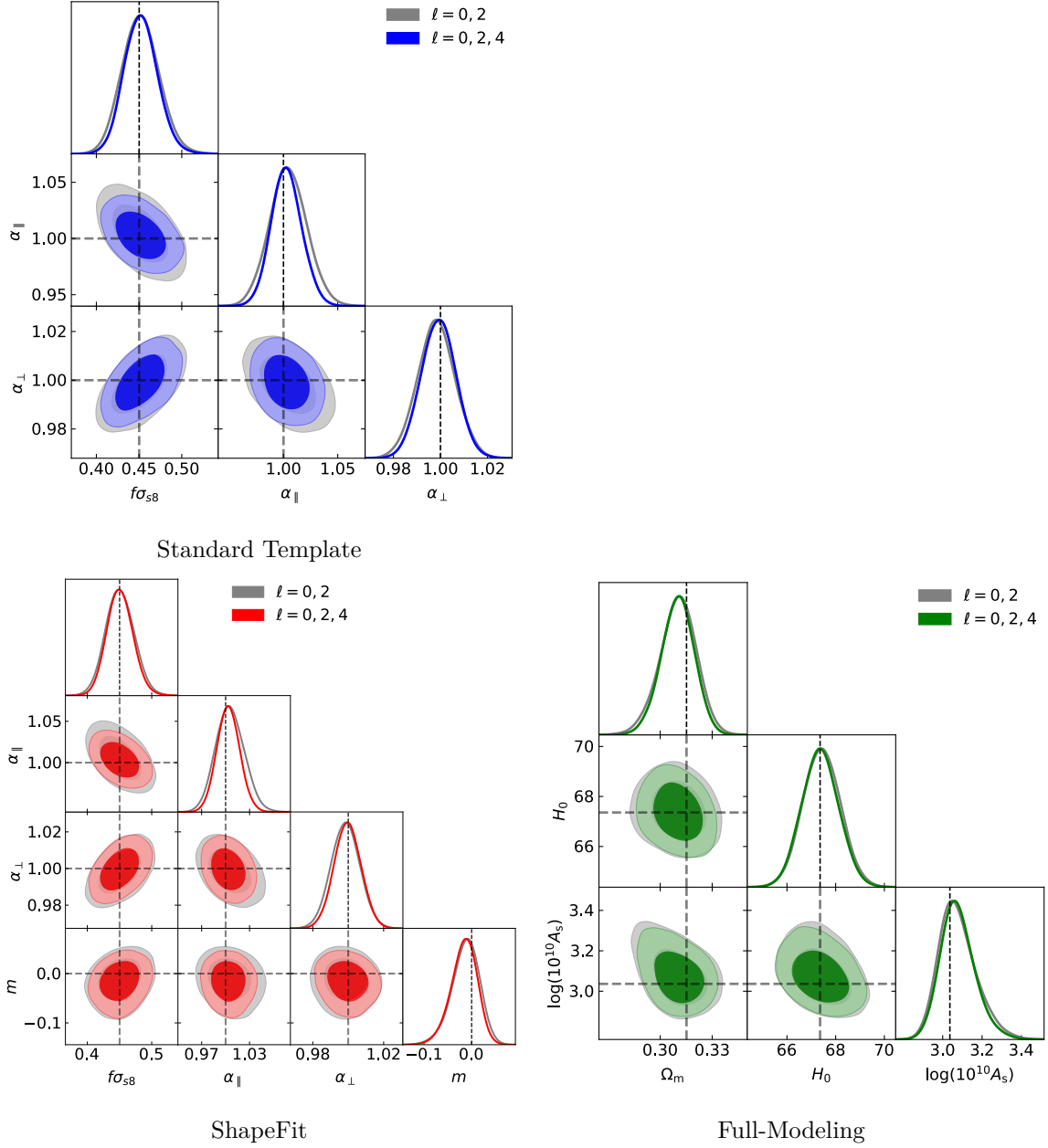
Figure 21. Comparison of constraints between $\ell = 0, 2$ and $\ell = 0, 2, 4$. We present fits using the covariance for both the single-box volume $(1 \cdot V)$.

E   Emulator error/performance

In order to speed up likelihood evaluations, we employ emulators that reproduce the theoretical power spectrum multipole predictions using a Taylor series centered on reasonably chosen values for the cosmological parameters, $\mathbf{\Omega}_0$, i.e. the Abacus fiducial values. The emulator is trained by evaluating the full `velocileptors` prediction on a grid with 9 points in each parameter direction, resulting in 9^N evaluations for N cosmological parameters. For each training point, e.g. $\mathbf{\Omega}_n = (h, \omega_b, \omega_{\mathrm{cdm}}, \log(10^{10} A_s))_n$ `velocileptors` computes the power
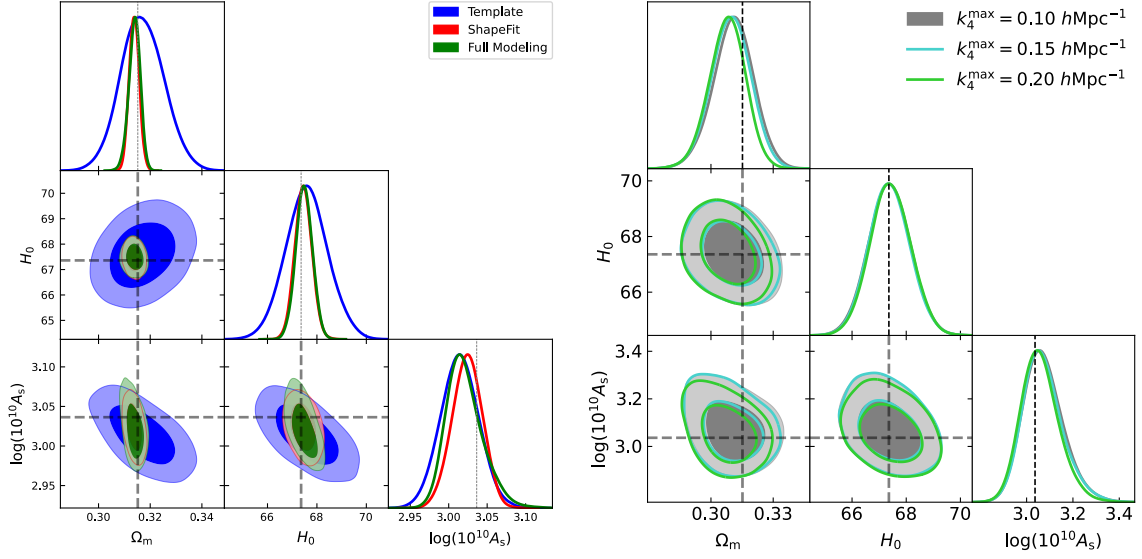
Figure 22. *(left):* constraints on ΛCDM parameters for the three modeling methods with hexadecapole included. *(right):* comparison of constraints on ΛCDM parameters, varying k_{max} of the hexadecapole while keeping k_{max} of the monopole and quadrupole at 0.2 $h\mathrm{Mpc}^{-1}$.

spectrum multipoles and separates the 19 terms within each multipole (i.e. the terms $P_{\ell,m}$ multiplied by $1, b_1, b_1^2, b_1 b_2$, etc.) into a table. After the grid of $P_{\ell,m}(\mathbf{\Omega}_n, k)$ has been computed for every n'th set of cosmological parameters, we take numerical derivatives up to fourth order in each parameter using the finite differencing method.[19] These arrays of derivatives are then stored for later use. At each step of an MCMC, the emulated power spectrum multipole terms are produced for the proposal set of parameters $\mathbf{\Omega}_n$ by constructing the Taylor series:

$$P_{\ell,m}^{\mathrm{emu}}(\mathbf{\Omega}_n, k) = P_{\ell,m}(\mathbf{\Omega}_0, k) + \sum_i^N \frac{\partial P_{\ell,m}}{\partial \mathbf{\Omega}_i}(\mathbf{\Omega}_{0,i} - \mathbf{\Omega}_{n,i}) +$$

$$+ \frac{1}{2} \sum_{i,j}^N \frac{\partial^2 P_{\ell,m}}{\partial \mathbf{\Omega}_i \partial \mathbf{\Omega}_j}(\mathbf{\Omega}_{0,i} - \mathbf{\Omega}_{n,i})(\mathbf{\Omega}_{0,j} - \mathbf{\Omega}_{n,j}) + \ldots \quad \text{(E.1)}$$

where $\mathbf{\Omega}_0$ is the set of cosmological parameters that the Taylor series was centered around, $\mathbf{\Omega}_{0,i}$ is the i'th cosmological parameter in said vector, and N is the number of parameters being varied in $\mathbf{\Omega}$. In order to demonstrate the accuracy of the emulator, we perform fits to the LRG cubic mocks both with the emulator and without. The results are shown in figure 23 for ShapeFit and Full-Modeling. In both cases, the emulator reproduces the constraints of the direct computation exactly.

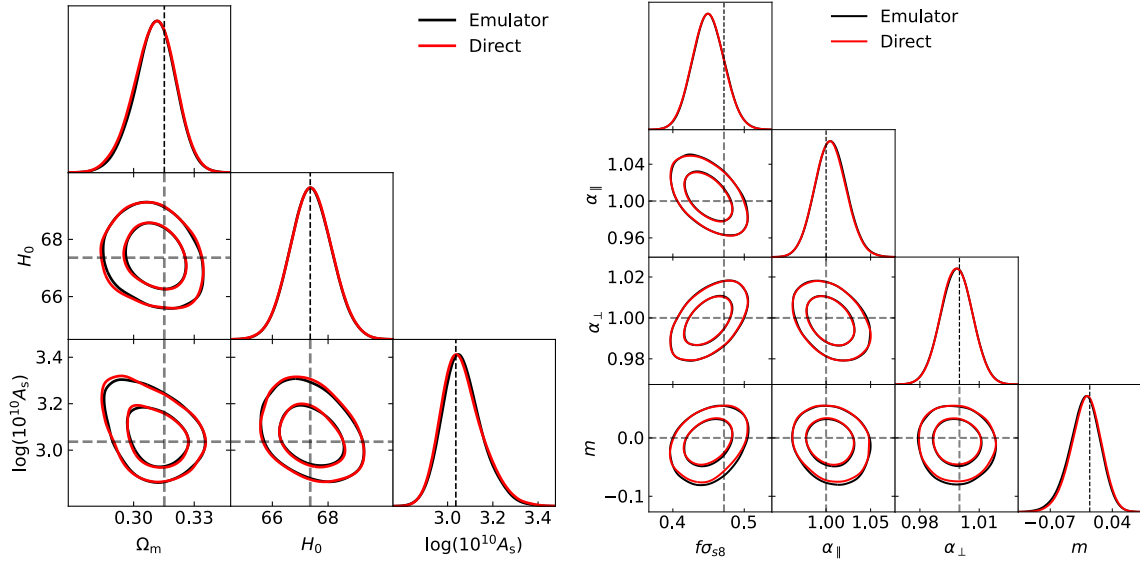[19]`findiff`; https://github.com/maroba/findiff [101].

Figure 23. Comparison of constraints from fitting LRG Cubic mock data using the Taylor series emulator vs. a direct fit in Full-Modeling *(left)* and ShapeFit *(right)*.

References

[1] P.J. Peebles, *The Large-Scale Structure of the Universe*, Princeton University Press (1980) [INSPIRE].

[2] J.A. Peacock, *Cosmological physics*, Cambridge University Press (1999) [`DOI:10.1017/CBO9780511804533`] [INSPIRE].

[3] S. Dodelson, *Modern Cosmology*, Academic Press (2003) [INSPIRE].

[4] D. Baumann, *Cosmology*, Cambridge University Press (2022) [`DOI:10.1017/9781108937092`] [INSPIRE].

[5] 2DFGRS collaboration, *The 2dF Galaxy Redshift Survey: Spectra and redshifts*, *Mon. Not. Roy. Astron. Soc.* **328** (2001) 1039 [`astro-ph/0106498`] [INSPIRE].

[6] D.H. Jones et al., *The 6dF Galaxy Survey: Final Redshift Release (DR3) and Southern Large-Scale Structures*, *Mon. Not. Roy. Astron. Soc.* **399** (2009) 683 [`arXiv:0903.5451`] [INSPIRE].

[7] S.P. Driver et al., *GAMA: towards a physical understanding of galaxy formation*, *Astron. Geophys.* **50** (2009) 5.12 [`arXiv:0910.5123`] [INSPIRE].

[8] M.J. Drinkwater et al., *The WiggleZ Dark Energy Survey: Survey Design and First Data Release*, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 1429 [`arXiv:0911.4246`] [INSPIRE].

[9] SDSS collaboration, *The Sloan Digital Sky Survey: Technical Summary*, *Astron. J.* **120** (2000) 1579 [`astro-ph/0006396`] [INSPIRE].

[10] BOSS collaboration, *SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way Galaxy, and Extra-Solar Planetary Systems*, *Astron. J.* **142** (2011) 72 [`arXiv:1101.1529`] [INSPIRE].

[11] BOSS collaboration, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2617 [arXiv:1607.03155] [INSPIRE].

[12] BOSS collaboration, *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III*, *Astrophys. J. Suppl.* **219** (2015) 12 [arXiv:1501.00963] [INSPIRE].

[13] EBOSS collaboration, *The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations with Lyα Forests*, *Astrophys. J.* **901** (2020) 153 [arXiv:2007.08995] [INSPIRE].

[14] EBOSS collaboration, *The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Large-scale Structure Catalogues and Measurement of the isotropic BAO between redshift 0.6 and 1.1 for the Emission Line Galaxy Sample*, *Mon. Not. Roy. Astron. Soc.* **500** (2020) 3254 [arXiv:2007.09007] [INSPIRE].

[15] EBOSS collaboration, *The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release*, *Astrophys. J. Suppl.* **250** (2020) 8 [arXiv:2007.09001] [INSPIRE].

[16] EUCLID collaboration, *Euclid Definition Study Report*, arXiv:1110.3193 [INSPIRE].

[17] L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **21** (2018) 2 [arXiv:1606.00180] [INSPIRE].

[18] DESI collaboration, *The DESI Experiment Part I: Science,Targeting, and Survey Design*, arXiv:1611.00036 [INSPIRE].

[19] DESI collaboration, *The DESI Experiment Part II: Instrument Design*, arXiv:1611.00037 [INSPIRE].

[20] DESI collaboration, *Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **164** (2022) 207 [arXiv:2205.10939] [INSPIRE].

[21] DESI collaboration, *Validation of the Scientific Program for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **167** (2024) 62 [arXiv:2306.06307] [INSPIRE].

[22] DESI collaboration, *The Early Data Release of the Dark Energy Spectroscopic Instrument*, *Astron. J.* **168** (2024) 58 [arXiv:2306.06308] [INSPIRE].

[23] DESI collaboration, *DESI 2024 I: Data Release 1 of the Dark Energy Spectroscopic Instrument*, in preparation (2025).

[24] DESI collaboration, *DESI 2024 II: Sample definitions, characteristics and two-point clustering statistics*, in preparation (2024).

[25] DESI collaboration, *DESI 2024 III: Baryon Acoustic Oscillations from Galaxies and Quasars*, arXiv:2404.03000 [INSPIRE].

[26] DESI collaboration, *DESI 2024 V: Analysis of the full shape of two-point clustering statistics from galaxies and quasars*, in preparation (2024).

[27] DESI collaboration, *DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman Alpha Forest*, arXiv:2404.03001 [INSPIRE].

[28] DESI collaboration, *DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations*, arXiv:2404.03002 [INSPIRE].

[29] DESI collaboration, *DESI 2024 VII: Cosmological constraints from full-shape analyses of the two-point clustering statistics measurements*, in preparation (2024).

[30] DESI collaboration, *DESI 2024 VIII: Constraints on Primordial Non-Gaussianities*, in preparation (2024).

[31] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rept.* **733** (2018) 1 [arXiv:1611.09787] [InSPIRE].

[32] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1 [InSPIRE].

[33] A.J.S. Hamilton, *Measuring Omega and the real correlation function from the redshift correlation function*, *Astrophys. J. Lett.* **385** (1992) L5 [InSPIRE].

[34] J.J.M. Carrasco, M.P. Hertzberg and L. Senatore, *The Effective Field Theory of Cosmological Large Scale Structures*, *JHEP* **09** (2012) 082 [arXiv:1206.2926] [InSPIRE].

[35] R.A. Porto, L. Senatore and M. Zaldarriaga, *The Lagrangian-space Effective Field Theory of Large Scale Structures*, *JCAP* **05** (2014) 022 [arXiv:1311.2168] [InSPIRE].

[36] Z. Vlah, M. White and A. Aviles, *A Lagrangian effective field theory*, *JCAP* **09** (2015) 014 [arXiv:1506.05264] [InSPIRE].

[37] S.-F. Chen, Z. Vlah and M. White, *Consistent Modeling of Velocity Statistics and Redshift-Space Distortions in One-Loop Perturbation Theory*, *JCAP* **07** (2020) 062 [arXiv:2005.00523] [InSPIRE].

[38] S.-F. Chen, Z. Vlah, E. Castorina and M. White, *Redshift-Space Distortions in Lagrangian Perturbation Theory*, *JCAP* **03** (2021) 100 [arXiv:2012.04636] [InSPIRE].

[39] G. D'Amico, L. Senatore and P. Zhang, *Limits on wCDM from the EFTofLSS with the PyBird code*, *JCAP* **01** (2021) 006 [arXiv:2003.07956] [InSPIRE].

[40] G. D'Amico et al., *The Cosmological Analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure*, *JCAP* **05** (2020) 005 [arXiv:1909.05271] [InSPIRE].

[41] T. Colas et al., *Efficient Cosmological Analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure*, *JCAP* **06** (2020) 001 [arXiv:1909.07951] [InSPIRE].

[42] H.E. Noriega, A. Aviles, S. Fromenteau and M. Vargas-Magaña, *Fast computation of non-linear power spectrum in cosmologies with massive neutrinos*, *JCAP* **11** (2022) 038 [arXiv:2208.02791] [InSPIRE].

[43] S. Ramirez et al., *Full Shape Cosmology Analysis from BOSS in configuration space using Neural Network Acceleration*, arXiv:2310.17834 [InSPIRE].

[44] H.E. Noriega et al., *Comparing Compressed and Full-modeling Analyses with FOLPS: Implications for DESI 2024 and beyond*, arXiv:2404.07269 [InSPIRE].

[45] Y. Lai et al., *A comparison between Shapefit compression and Full-Modelling method with PyBird for DESI 2024 and beyond*, arXiv:2404.07283 [InSPIRE].

[46] DESI collaboration, *Full Modeling and Parameter Compression Methods in configuration space for DESI 2024 and beyond*, arXiv:2404.07268 [InSPIRE].

[47] M. Maus et al., *A comparison of effective field theory models of redshift space galaxy power spectra for DESI 2024 and future surveys*, arXiv:2404.07272 [InSPIRE].

[48] S.-F. Chen, Z. Vlah and M. White, *A new analysis of galaxy 2-point functions in the BOSS survey, including full-shape information and post-reconstruction BAO*, *JCAP* **02** (2022) 008 [arXiv:2110.05530] [InSPIRE].

[49] S.-F. Chen, Z. Vlah and M. White, *The reconstructed power spectrum in the Zeldovich approximation*, *JCAP* **09** (2019) 017 [arXiv:1907.00043] [InSPIRE].

[50] 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].

[51] S. Brieden, H. Gil-Marín and L. Verde, *ShapeFit: extracting the power spectrum shape information in galaxy surveys beyond BAO and RSD*, *JCAP* **12** (2021) 054 [arXiv:2106.07641] [InSPIRE].

[52] M. Maus, S.-F. Chen and M. White, *A comparison of template vs. direct model fitting for redshift-space distortions in BOSS*, *JCAP* **06** (2023) 005 [arXiv:2302.07430] [InSPIRE].

[53] L.H. Garrison et al., *The abacus cosmological N-body code*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 575 [arXiv:2110.11392] [InSPIRE].

[54] C.-H. Chuang et al., *EZmocks: extending the Zel'dovich approximation to generate mock galaxy catalogues with accurate clustering statistics*, *Mon. Not. Roy. Astron. Soc.* **446** (2015) 2621 [arXiv:1409.1124] [InSPIRE].

[55] C. Grove et al., *The DESI N-body simulation project — I. Testing the robustness of simulations for the DESI dark time survey*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 1854 [arXiv:2112.09138] [InSPIRE].

[56] R.E. Angulo and O. Hahn, *Large-scale dark matter simulations*, arXiv:2112.05165 [DOI:10.1007/s41115-021-00013-z] [InSPIRE].

[57] J. Hartlap, P. Simon and P. Schneider, *Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix*, *Astron. Astrophys.* **464** (2007) 399 [astro-ph/0608064] [InSPIRE].

[58] M.M. Abidi and T. Baldauf, *Cubic Halo Bias in Eulerian and Lagrangian Space*, *JCAP* **07** (2018) 029 [arXiv:1802.07622] [InSPIRE].

[59] J. Carlson, B. Reid and M. White, *Convolution Lagrangian perturbation theory for biased tracers*, *Mon. Not. Roy. Astron. Soc.* **429** (2013) 1674 [arXiv:1209.0780] [InSPIRE].

[60] L. Senatore and M. Zaldarriaga, *The IR-resummed Effective Field Theory of Large Scale Structures*, *JCAP* **02** (2015) 013 [arXiv:1404.5954] [InSPIRE].

[61] D. Blas, M. Garny, M.M. Ivanov and S. Sibiryakov, *Time-Sliced Perturbation Theory II: Baryon Acoustic Oscillations and Infrared Resummation*, *JCAP* **07** (2016) 028 [arXiv:1605.02149] [InSPIRE].

[62] Z. Vlah, U. Seljak, M.Y. Chu and Y. Feng, *Perturbation theory, effective field theory, and oscillations in the power spectrum*, *JCAP* **03** (2016) 057 [arXiv:1509.02120] [InSPIRE].

[63] Z. Vlah, E. Castorina and M. White, *The Gaussian streaming model and convolution Lagrangian effective field theory*, *JCAP* **12** (2016) 007 [arXiv:1609.02908] [InSPIRE].

[64] M. Schmittfull et al., *Modeling Galaxies in Redshift Space at the Field Level*, *JCAP* **05** (2021) 059 [arXiv:2012.03334] [InSPIRE].

[65] C. Alcock and B. Paczynski, *An evolution free test for non-zero cosmological constant*, *Nature* **281** (1979) 358 [InSPIRE].

[66] N. Findlay et al., *Fiducial cosmology impact for DESI 2024 full shape analysis*, in preparation (2024).

[67] R.J. Cooke, M. Pettini and C.C. Steidel, *One Percent Determination of the Primordial Deuterium Abundance*, *Astrophys. J.* **855** (2018) 102 [arXiv:1710.11129] [InSPIRE].

[68] S. Brieden, H. Gil-Marín and L. Verde, *A tale of two (or more) h's*, *JCAP* **04** (2023) 023 [arXiv:2212.04522] [INSPIRE].

[69] D.J. Eisenstein, H.-J. Seo, E. Sirko and D. Spergel, *Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak*, *Astrophys. J.* **664** (2007) 675 [astro-ph/0604362] [INSPIRE].

[70] Y. Noh, M. White and N. Padmanabhan, *Reconstructing baryon oscillations*, *Phys. Rev. D* **80** (2009) 123501 [arXiv:0909.1802] [INSPIRE].

[71] N. Padmanabhan, M. White and J.D. Cohn, *Reconstructing Baryon Oscillations: A Lagrangian Theory Perspective*, *Phys. Rev. D* **79** (2009) 063523 [arXiv:0812.2905] [INSPIRE].

[72] M. White, *Reconstruction within the Zeldovich approximation*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 3822 [arXiv:1504.03677] [INSPIRE].

[73] S.-F. Chen et al., *Baryon Acoustic Oscillation Theory and Modelling Systematics for the DESI 2024 results*, arXiv:2402.14070 [INSPIRE].

[74] B. Wallisch, *Cosmological Probes of Light Relics*, Ph.D. Thesis, University of Cambridge, U.K. (2018) [arXiv:1810.02800] [INSPIRE].

[75] N. Sugiyama, *Developing a Theoretical Model for the Resummation of Infrared Effects in the Post-Reconstruction Power Spectrum (an explanatory video is available in the comments section)*, arXiv:2402.06142 [INSPIRE].

[76] R.W. Hockney and J.W. Eastwood, *Computer simulation using particles*, CRC Press (1988) [DOI:10.1201/9780367806934].

[77] D. Jeong, *Cosmology with high ($z > 1$) redshift galaxy surveys*, Ph.D. Thesis, University of Texas, Austin, U.S.A. (2010).

[78] PLANCK collaboration, *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, *Astron. Astrophys.* **641** (2020) A1 [arXiv:1807.06205] [INSPIRE].

[79] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [arXiv:1807.06209] [INSPIRE].

[80] 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].

[81] S.L. Bridle et al., *Analytic marginalization over CMB calibration and beam uncertainty*, *Mon. Not. Roy. Astron. Soc.* **335** (2002) 1193 [astro-ph/0112114] [INSPIRE].

[82] A.N. Taylor and T.D. Kitching, *Analytic Methods for Cosmological Likelihoods*, *Mon. Not. Roy. Astron. Soc.* **408** (2010) 865 [arXiv:1003.1136] [INSPIRE].

[83] W. Handley and P. Lemos, *Quantifying tensions in cosmological parameters: Interpreting the DES evidence ratio*, *Phys. Rev. D* **100** (2019) 043504 [arXiv:1902.04029] [INSPIRE].

[84] DES collaboration, *Assessing tension metrics with dark energy survey and Planck data*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 6179 [arXiv:2012.09554] [INSPIRE].

[85] A. Gómez-Valent, *Fast test to assess the impact of marginalization in Monte Carlo analyses and its application to cosmology*, *Phys. Rev. D* **106** (2022) 063506 [arXiv:2203.16285] [INSPIRE].

[86] B. Hadzhiyska et al., *Cosmology with 6 parameters in the Stage-IV era: efficient marginalisation over nuisance parameters*, arXiv:2301.11895 [DOI:10.21105/astro.2301.11895] [INSPIRE].

[87] N. Sailer et al., *Cosmological constraints from the cross-correlation of DESI Luminous Red Galaxies with CMB lensing from Planck PR4 and ACT DR6*, arXiv:2407.04607 [INSPIRE].

[88] S.-F. Chen, M. White, J. DeRose and N. Kokron, *Cosmological analysis of three-dimensional BOSS galaxy clustering and Planck CMB lensing cross correlations via Lagrangian perturbation theory*, *JCAP* **07** (2022) 041 [arXiv:2204.10392] [INSPIRE].

[89] A. Syversveen, *Noninformative bayesian priors. interpretation and problems with construction and applications*, https://www.researchgate.net/publication/2288953_Noninformative_Bayesian_Priors_Interpretation_And_Problems_With_Construction_And_Applications (1998).

[90] J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*, *Meas. Sci. Technol.* **12** (2001) 221.

[91] U. Seljak, *Analytic model for galaxy and dark matter clustering*, *Mon. Not. Roy. Astron. Soc.* **318** (2000) 203 [astro-ph/0001493] [INSPIRE].

[92] J.A. Peacock and R.E. Smith, *Halo occupation numbers and galaxy bias*, *Mon. Not. Roy. Astron. Soc.* **318** (2000) 1144 [astro-ph/0005010] [INSPIRE].

[93] E. Schaan and M. White, *Multi-tracer intensity mapping: Cross-correlations, Line noise & Decorrelation*, *JCAP* **05** (2021) 068 [arXiv:2103.01964] [INSPIRE].

[94] R. Scoccimarro, *Redshift-space distortions, pairwise velocities and nonlinearities*, *Phys. Rev. D* **70** (2004) 083007 [astro-ph/0407214] [INSPIRE].

[95] Z. Vlah and M. White, *Exploring redshift-space distortions in large-scale structure*, *JCAP* **03** (2019) 007 [arXiv:1812.02775] [INSPIRE].

[96] M.M. Ivanov and O.H.E. Philcox, *Measuring H_0 with Spectroscopic Surveys*, arXiv:2305.07977 [INSPIRE].

[97] M.M. Ivanov, M. Simonović and M. Zaldarriaga, *Cosmological Parameters from the BOSS Galaxy Power Spectrum*, *JCAP* **05** (2020) 042 [arXiv:1909.05277] [INSPIRE].

[98] BOSS collaboration, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Testing gravity with redshift-space distortions using the power spectrum multipoles*, *Mon. Not. Roy. Astron. Soc.* **443** (2014) 1065 [arXiv:1312.4611] [INSPIRE].

[99] K.C. Chan, R. Scoccimarro and R.K. Sheth, *Gravity and Large-Scale Non-local Bias*, *Phys. Rev. D* **85** (2012) 083509 [arXiv:1201.3614] [INSPIRE].

[100] N. Kokron et al., *The cosmology dependence of galaxy clustering and lensing from a hybrid N-body-perturbation theory model*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 1422 [arXiv:2101.11014] [INSPIRE].

[101] M. Baer, *findiff software package*, https://github.com/maroba/findiff (2018).

Author List

M. Maus ^{a,b,*}, S. Chen ^{c}, M. White ^{a,b,d}, J. Aguilar ^{b}, S. Ahlen ^{e}, A. Aviles ^{f,g}, S. Brieden ^{h},
D. Brooks ^{i}, T. Claybaugh ^{b}, S. Cole ^{j}, A. de la Macorra ^{k}, Arjun Dey ^{l}, P. Doel ^{i}, S. Ferraro ^{b,m},
N. Findlay ^{n}, J. E. Forero-Romero ^{o,p}, E. Gaztañaga ^{q,n,r}, H. Gil-Marín ^{s,q,t},
S. Gontcho A Gontcho ^{b}, C. Hahn ^{u}, K. Honscheid ^{v,w,x}, C. Howlett ^{y}, M. Ishak ^{z}, S. Juneau ^{l},
A. Kremin ^{b}, Y. Lai ^{y}, M. Landriau ^{b}, M. E. Levi ^{b}, M. Manera ^{aa,ab}, R. Miquel ^{ac,ab},
E. Mueller ^{ad}, A. D. Myers ^{ae}, S. Nadathur ^{n}, J. Nie ^{af}, H. E. Noriega ^{g,k},
N. Palanque-Delabrouille ^{ag,b}, W. J. Percival ^{ah,ai,aj}, C. Poppett ^{b,ak,m}, S. Ramirez-Solano ^{k},
M. Rezaie ^{al}, A. Rocher ^{am,ag}, G. Rossi ^{an}, E. Sanchez ^{ao}, D. Schlegel ^{b}, M. Schubnell ^{ap,aq},
H. Seo ^{ar}, D. Sprayberry ^{l}, G. Tarlé ^{aq}, M. Vargas-Magaña ^{k}, B. A. Weaver ^{l}, S. Yuan ^{as},
P. Zarrouk ^{at}, H. Zhang ^{ah,aj}, R. Zhou ^{b}, H. Zou ^{af}

^{a} Department of Physics, University of California, Berkeley, CA 94720, U.S.A.
^{b} Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.
^{c} Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, U.S.A.
^{d} Department of Astronomy, University of California at Berkeley,
   501 Campbell Hall # 3411, Berkeley CA, U.S.A.
^{e} Physics Department, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, U.S.A.
^{f} Instituto Avanzado de Cosmología A. C.,
   San Marcos 11 – Atenas 202. Magdalena Contreras, 10720. Ciudad de México, México
^{g} Instituto de Ciencias Físicas, Universidad Autónoma de México,
   Cuernavaca, Morelos, 62210, México
^{h} Institute for Astronomy, University of Edinburgh, Royal Observatory,
   Blackford Hill, Edinburgh EH9 3HJ, U.K.
^{i} Department of Physics & Astronomy, University College London,
   Gower Street, London, WC1E 6BT, U.K.
^{j} Institute for Computational Cosmology, Department of Physics, Durham University,
   South Road, Durham DH1 3LE, U.K.
^{k} Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México
^{l} NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, U.S.A.
^{m} University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, U.S.A.
^{n} Institute of Cosmology and Gravitation, University of Portsmouth,
   Dennis Sciama Building, Portsmouth, PO1 3FX, U.K.
^{o} Departamento de Física, Universidad de los Andes,
   Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia
^{p} Observatorio Astronómico, Universidad de los Andes,
   Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia
^{q} Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
^{r} Institute of Space Sciences, ICE-CSIC, Campus UAB,
   Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain
^{s} Departament de Física Quàntica i Astrofísica, Universitat de Barcelona,
   Martí i Franquès 1, E08028 Barcelona, Spain
^{t} Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB),
   c. Martí i Franquès, 1, 08028 Barcelona, Spain
^{u} Department of Astrophysical Sciences, Princeton University, Princeton NJ 08544, U.S.A.
^{v} Center for Cosmology and AstroParticle Physics, The Ohio State University,
   191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.
^{w} Department of Physics, The Ohio State University,
   191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.
^{x} The Ohio State University, Columbus, 43210 OH, U.S.A.
^{y} School of Mathematics and Physics, University of Queensland, 4072, Australia

^{z} *Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, U.S.A.*

^{aa} *Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona,*
08193 Bellaterra (Barcelona), Spain

^{ab} *Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology,*
Campus UAB, 08193 Bellaterra Barcelona, Spain

^{ac} *Institució Catalana de Recerca i Estudis Avançats,*
Passeig de Lluís Companys, 23, 08010 Barcelona, Spain

^{ad} *Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, U.K.*

^{ae} *Department of Physics & Astronomy, University of Wyoming,*
1000 E. University, Dept. 3905, Laramie, WY 82071, U.S.A.

^{af} *National Astronomical Observatories, Chinese Academy of Sciences,*
A20 Datun Rd., Chaoyang District, Beijing, 100012, P.R. China

^{ag} *IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*

^{ah} *Department of Physics and Astronomy, University of Waterloo,*
200 University Ave W, Waterloo, ON N2L 3G1, Canada

^{ai} *Perimeter Institute for Theoretical Physics,*
31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada

^{aj} *Waterloo Centre for Astrophysics, University of Waterloo,*
200 University Ave W, Waterloo, ON N2L 3G1, Canada

^{ak} *Space Sciences Laboratory, University of California, Berkeley,*
7 Gauss Way, Berkeley, CA 94720, U.S.A.

^{al} *Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, U.S.A.*

^{am} *Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

^{an} *Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea*

^{ao} *CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*

^{ap} *Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A.*

^{aq} *University of Michigan, Ann Arbor, MI 48109, U.S.A.*

^{ar} *Department of Physics & Astronomy, Ohio University, Athens, OH 45701, U.S.A.*

^{as} *SLAC National Accelerator Laboratory, Menlo Park, CA 94305, U.S.A.*

^{at} *Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies*
(LPNHE), FR-75005 Paris, France

* *Corresponding author*